Ч. Кантор, П. Шиммел

БИОФИЗИЧЕСКАЯ ХИМИЯ

ИЗДАТЕЛЬСТВО МИР

BIOPHYSICAL CHEMISTRY

Part I

The Conformation of Biological Macromolecules

Charles R. Cantor
Columbia University

Paul R. Schimmel
Massachusetts Institute of Technology

W.H.Freeman and Company San Francisco

Ч.Кантор, П.Шиммел

БИОФИЗИЧЕСКАЯ ХИМИЯ

В трех томах Том 1

Перевод с английского

Под редакцией
д-ра хим. наук А.А.Богданова,
д-ра физ.-мат. наук Ю.С.Лазуркина,
д-ра физ.-мат. наук М.Д.Франк-Каменецкого

МОСКВА «МИР» 1984

ББК 28.072
К19
УДК 577

Кантор Ч., Шиммел П.

К19 Биофизическая химия: В 3-х т. Пер. с англ. — М.: Мир, 1984. — Т.1 — 336 с., ил.

В трехтомном издании, написанном ученымн нз США, на самом современном уровне нзложены основные представления о биологическнх макромолекулах и методах исследования нх структуры н функций. В первом томе рассмотрены общие принципы организации первичной, вторичной, третичной и четвертнчной структуры белков н нуклеиновых кислот, а также строение полисахаридов, нуклеопротеинов и биологических мембран. Книга написана ясно и четко, на очень высоком научном уровне.

Для биофнзиков, биохимиков, молекулярных бнологов, физиков, химиков, для преподавателей, аспирантов н студентов бнологических специальностей.

К $\frac{2001040000 - 171}{041(01)-84}$ св. пл. подп. изд-ий, 83

ББК 28.072
57.04

Редакция литературы по биологии

Чарлз Р. Кантор, Пол Р. Шиммел
БИОФИЗИЧЕСКАЯ ХИМИЯ
том 1

Научные редакторы Н. А. Райская, Н. Н. Шафрановская.
Мл. редактор З. Е. Кожанова. Художник Б. А. Шляпугин.
Художественный редактор Ю. Л. Максимов. Технический редактор Л. П. Чуркина.
Корректоры Р. Л. Вибке, И. А. Шурша

ИБ № 3071

Подписано к печати 29.03.84 г. Формат 70 × 100¹/₁₆. Бумага офсетная № 1.
Гарнитура Таймс. Печать офсетная. Объем 10,5 бум. л. Усл. печ. л. 27,3.
Усл. кр.-отт. 41,00. Уч.-изд. л. 27,87. Изд. № 4/1930. Тираж 8300 экз.
Зак. 1141. Цена 2 р. 50 к.

Набрано в издательстве «Мир» на фотонаборном комплексе «Компьюграфик»
129820, ГСП Москва, 1-й Рижский пер., 2.

Можайский полиграфкомбинат Союзполиграфпрома при Государственном
комитете СССР по делам издательств, полиграфии и книжной торговли.
143200, Можайск, ул. Мира, 93.

© 1980 by W. H. Freeman and Company

© Перевод на русский язык,
«Мир», 1984

Предисловие редакторов перевода

Тот бурный прогресс в понимании основ жизненных процессов, свидетелями которого мы являемся, есть результат необычайно плодотворного слияния трех наук — физики, химии и биологии. Сформировавшуюся в результате этого слияния синтетическую дисциплину называют по-разному — и физико-химическая биология, и молекулярная биофизика, и физическая биохимия, и биофизическая химия. Суть от этого не меняется — речь идет о применении физических и химических идей и методов к изучению биологических процессов на молекулярном и атомном уровнях.

Синтетический характер этой новой науки, взрывообразное ее развитие делают необычайно сложной задачу создания учебных пособий; в то же время в них ощущается крайне острая потребность. Вышедшие в последние годы отечественные и переводные книги такого рода (Волькенштейн М.В. Молекулярная биофизика. — М.: Наука, 1975; Волькенштейн М.В. Биофизика. — М.: Наука, 1981; Фрайфельдер Д. Физическая биохимия. Пер. с англ. — М.: Мир, 1980; Маршелл Э. Биофизическая химия. Пер. с англ. — М.: Мир, 1981, и другие) лишь частично решали проблему. Предлагаемый вниманию читателей трехтомник является, на наш взгляд, наиболее удачной попыткой создания последовательного курса в данной области. Его авторы, Ч.Кантор и П.Шиммел, — известные американские ученые, занимающиеся вопросами строения и функционирования комплексов белков и нуклеиновых кислот. В течение многих лет они читают лекции по молекулярной биологии и биофизике в крупнейших высших учебных заведениях США — первый в Колумбийском университете, второй в Массачусетском технологическом институте. В результате огромной работы по отбору, систематизации, упорядочению материала авторам удалось создать законченный и очень полезный курс биофизической химии.

План построения монографии вполне естествен. Первый том вводит читателя в основные проблемы данной науки. В нем рассказывается о строении важнейших молекул живой клетки. Второй том чисто методический. Здесь представлены те физические методы, которые наиболее широко используются в биофизической химии. И, наконец, в третьем томе рассказано о свойствах биополимеров и о связи их структуры и свойств с некоторыми биологическими функциями. Материал в основном доступен начинающему читателю, впервые знакомящемуся с данной областью. Удачен подбор задач и многочисленных конкретных примеров, которые разобраны с необходимой полнотой.

Курс знакомит читателя с наиболее актуальными проблемами биофизической химии. Но эта область развивается такими темпами, что за период, прошедший со времени написания книги, до завершения перевода ее на русский язык появилось много новых работ принципиального характера, главным образом по темам, входящим в третий том. В связи в этим мы позволили себе дать небольшое число примечаний, содержащих ссылки на наиболее важные новые работы, в особениости обзоры.

Не вызывает сомнений, что это фундаментальное руководство будет очень полезным как студентам и аспирантам химических, биологических, биофизических специальностей, так и исследователям, работающим в области физико-химической биологии, молекулярной биофизики и в смежных областях.

А.А. Богданов
Ю.С. Лазуркин
М.Д. Франк-Каменецкий

Луису и Иде Диане Кантор,
Альфреду и Дорис Шиммел

Предисловие

Эта книга посвящена биологическим макромолекулам и их комплексам. В ней рассмотрены конформация молекул, их форма, структура, конформационные изменения и межмолекулярные взаимодействия, а также изложены основные принципы и закономерности, лежащие в основе биофизической химии и вытекающие из законов физики, химии и биологии.

Мы попытались написать учебное пособие, рассчитанное на читателей с разным уровнем подготовки. Материал сильно различается по степени сложности, так что учебник будет полезен для студентов и младших, и старших курсов. Он может представлять интерес и для научных работников — как неспециалистов в данной области, желающих получить о ней представление, так и опытных биофизиков.

В первом томе рассмотрены структура биологических макромолекул и силы, ее определяющие. Гл. 1 вводит в круг наиболее важных проблем биофизической химии. Гл. 2-4 знакомят со структурой белков, нуклеиновых кислот и других биополимеров, в гл. 5-6 рассмотрены нековалентные взаимодействия и конформационный анализ.

Второй том посвящен методам, использующимся при изучении структуры и функций биологических систем. При этом мы не ставили перед собой задачу описать все известные методы, подробно рассмотрены лишь наиболее важные из них. Гл. 7-9 посвящены спектроскопическим методам, гл. 10-12 — гидродинамическим, а в гл. 13 и 14 обсуждаются рентгеноструктурный анализ и другие дифракционные методы.

В третьем томе показано, как благодаря совместному использованию различных экспериментальных методов и теоретических подходов удается понять поведение и свойства биологических макромолекул. Главное внимание уделяется термодинамике и кинетике конформационных изменений и взаимодействию макромолекул с лигандами. При этом в случае необходимости описываются новые методы, а также подробно рассматривается история исследования некоторых вопросов. В гл. 15-17 обсуждаются проблемы взаимодействия молекул с лигандами, в гл. 18 и 19 излагаются теории и методы, используемые при изучении молекул, конформация которых носит статистический характер, гл. 20-24 посвящены конформационным изменениям в белках и нуклеиновых кислотах, гл. 25 — биологическим мембранам.

Мы постарались построить изложение так, чтобы главы можно было читать независимо одну от другой, тем самым предоставляя читателю возможность выборочного изучения материала. Этой же цели служат многочисленные перекрестные ссылки на вопросы, обсуждающиеся в разных главах. Всюду, где это возможно, в качестве примеров рассматриваются одни и те же системы, с тем чтобы читатель мог сопоставить сведения, получаемые относительно одного и того же белка или нуклеиновой кислоты с помощью различных подходов.

Мы стремились выдержать в пределах каждой главы примерно одинаковый уровень строгости и сложности изложения. Краткие отступления от этого уровня выделены в виде дополнений, более обширные помечены знаком ● при заголовке соответствующего раздела. В некоторых дополнениях кратко излагаются необходимые элементарные сведения из области математики, физики и физической химии. Они могут оказаться весьма полезными для читателя, не имеющего специальной подготовки по этим дисциплинам. Другие дополнения предназначены для более подготовленных читателей; во многих случаях

при этом обсуждаются вопросы, неясные для самих авторов. В приложении А приведены необходимые сведения из матричной алгебры.

Разделы книги сильно различаются по уровню сложности используемого математического аппарата. Всюду, где это не было в ущерб ясности изложения, мы старались обходиться как можно более простыми математическими выкладками. Например, гидродинамические свойства рассмотреиы только для одномерного случая. Ряд основных уравнений получен из анализа размерностей, а не путем долгих (и притом не особенно поучительных) выкладок, связаниых с рассмотрением граничных условий гидродинамических задач. С другой стороны, при обсуждении явления рассеяния рентгеновских лучей используется метод преобразования Фурье, а многие задачи статистической механики решаются методами матричной алгебры. Подобные математические методы используются лишь в некоторых главах; чтобы помочь читателю, ранее с ними не знакомому, мы посвятили этим вопросам многочисленные дополнения. Материал остальных разделов и глав может быть усвоен без привлечения современного математического аппарата.

Далеко не все методы и биологические объекты рассматриваются в книге сколько-нибудь подробно. Главным критерием при выборе была не важность того или другого метода, а то, насколько знакомство с ним будет полезно для начинающего изучать данную область.

В каждой главе имеются краткие выводы. Приведено много иллюстраций, в том числе рисунки Ирвинга Гейса, сделанные специально для этой книги. Одно только знакомство с краткими выводами и иллюстрациями позволит получить представление о предмете. Мы надеемся также, что иллюстрации помогут передать волнующую атмосферу, которая царит в этой области науки.

В конце каждой главы помещены задачи. Они сильно различаются по степени сложности. Среди них есть и сравнительно простые, и такие, для которых окончательный ответ неизвестен (по крайней мере авторам). Ответы к задачам даются в конце каждого тома.

В монографии нет подробных литературных ссылок, за исключением тех случаев, когда мы ссылаемся на источник, откуда был заимствован ранее опубликованный материал. Однако в каждой главе приводится список наиболее полезных работ. Практически во всех случаях этих ссылок достаточно, чтобы с их помощью найти литературу, необходимую для подробного изучения вопроса.

Трудной задачей оказалась унификация обозначений и сокращений. Используя материалы многочисленных исследований, различных по характеру, мы столкнулись с необходимостью унификации, исключающей неопределенности и ошибки при использовании сходных обозначений. Всюду, где это было возможно, мы следовали рекомендациям Американского химического общества, но в некоторых случаях нам пришлось решать вопрос на свой страх и риск. В конце третьего тома приведен список наиболее распространенных обозначений.

Часть материала книги была использована при чтении лекций по биофизической химии студентам старших курсов Массачусетского технологического института. Лекции предназначались для тех, кто хотел прослушать углубленный курс физической химии, с уклоном в сторону биохимии. Эти лекции охватывали в основном тот же материал и читались на том же уровне, что и традиционный курс физической химии. Предварительно студенты должны были прослушать односеместровый курс физической химии, который в МТИ посвящен преимущественно химической термодинамике.

В течение ряда лет материал, излагаемый в книге, использовался при чтении лекций для аспирантов в МТИ и в Колумбийском университете. Кроме того, часть материала входила в специальный курс, посвященный строению белков. Мы надеемся, что широта охвата сделает эту книгу полезной для преподавателей с самыми разными требованиями и инте-

ресами, а также позволит изменять содержание лекций в соответствии с развитием данной области.

Совершенно очевидно, что труд столь большой сложности не мог быть выполнен усилиями лишь двух авторов. В ходе написания книги, стремясь изложить и объяснить широкий круг проблем биофизической химии, мы поняли, почему за последние два десятилетия в этой области было написано так мало книг. Мы приносим глубокую благодарность всем, кто помог нам тем или иным путем — от разъяснения отдельных вопросов до предоставления оригинальных результатов.

Мы особенно благодарны Ирвингу Гейсу за создание ряда сложных иллюстраций и полезные советы относительно многих рисунков, а также Вильме Олсон, прочитавшей бóльшую часть рукописи, Роберту Олберти и Гордону Хэммсу за обсуждение вопросов, касающихся биохимического равновесия и кинетики, Ричарду Дикерсону за предоставление материалов и советы, оказавшиеся весьма полезными при подготовке гл.13, Полу Флори, вдохновившему нас на рассмотрение вопросов конформационной энергии и конфигурационной статистики макромолекул, Говарду Шахману, чей курс, читаемый в Беркли, побудил нас включить в монографию целый ряд разделов, Р.Вэйну Олеру, который свел авторов вместе для этой работы. Нам хотелось бы поблагодарить Брюса Армбрустера, заключившего с нами договор, и сотрудников издательства Фримен энд Компани, в том числе Рут Аллен, Артура Бартлетта, Роберта Иши, Ларри МакКомбса и Пэрл Вапнек. Мы весьма признательны Ким Энгель, Карен Хейнс, Мэри Людвиг, Джоан Мешне, Пегги Нелсон, Кэти Путлэнд и Джуди Шиммел за печатание рукописи и работу над ней. Нам хотелось бы поблагодарить Кассандру Смит и Джуди, Кэти и Кирстен Шиммел за то смирение, с которым они перенесли вторжение этого труда в жизнь авторов.

Многие люди прочитали отдельные главы и высказали критические замечания, предоставили нам иллюстративный материал и свои данные; они потратили много времени в полезных обсуждениях с нами. Это Роберт Олберти, Артур Арнон, Струтер Арнотт, П.Эткинс, Роберт Болдуин, Ларри Берлинер, Брюс Берн, Ричард Берзон, Шерман Бейхок, Виктор Блумфилд, Дэвид Брандт, Джои Брандтс, Джон Чамберс, Санни Чан, Патриция Коул, Роберт Кричтон, Фрэнсис Крик, Доналд Крозерс, Норман Дэвидсон, Ричард Дикерсон, Дэвид Эйзенберг, Роберт Фэйрклоф, Джерри Фасман, Джордж Флинн, Дэвид Фрайфелдер, Роналд Гэмбл, Роберт Дженнис, Маррей Гудмэн, Джонатан Грир, О.Хейс Гриффит, Гордон Хеммс, Джон Херст, Эллен Гендерсон, Джеймс Гилдебрандт, Рей Хестис, Сунг Хоу Ким, Арон Клуг, Нелсон Леонард, Х.Дж.Ли, Стефен Липпард, Ричард Лорд, Брайан Мэттьюз, Харден МакКоннел, Питер Мур, Гарт Николсон, Леонард Пеллер, Ричард Перхэм, Майкл Рафтери, Александр Рич, Фредерик Ричардс, Дэвид Ричардсон, Волфрам Сенджер, Говард Шахман, Харолд Шерага, Бенно Шенборн, Верн Шумейкер, Надриан Симэн, Роберт Шульман, Мэвис Шур, Луиза Слейд, Кассандра Смит, Хэнк Собелл, Томас Стейц, Роберт Строуд, Луберт Страйер, Серж Тимашев, Игнацио Тиноко мл., Ричард Вэндлен, Джером Виноград, Питер фон Хиппель, Кристофер Уолш, Джеймс Уонг, Грегорио Вебер, Питер Веллауэр, Барбара Уэллз, Роберт Уэллз, Уильм Уинтер, Гаролд Уайкоф, Джеффрис Уайман, Бруно Зимм. Мы выражаем всем им огромную благодарность.

Ноябрь 1979 г.

Чарлз Р.Кантор
Пол Р.Шиммел

Том 1

КОНФОРМАЦИЯ БИОЛОГИЧЕСКИХ МАКРОМОЛЕКУЛ

Эта книга посвящена применению физической химии для исследования свойств больших молекул, представляющих биологический интерес. Основное внимание в ней уделено белкам и нуклеиновым кислотам. Другие биополимеры, например полисахариды и липиды, рассмотрены менее подробно. Это связано с двумя обстоятельствами. Во-первых, белки и нуклеиновые кислоты играют главную роль во всех известных жизненных процессах. Во-вторых, их структура и функции в настоящее время изучены значительно лучше, чем аналогичные характеристики других биополимеров.

Первая часть книги решает несколько задач. Она вводит читателя в курс представлений о различных уровнях структурной организации биополимеров. В ней формулируются основные вопросы, на которые должен ответить исследователь, работающий в области физической химии биологических макромолекул. Высказываются также некоторые соображения об общих подходах к решению этих вопросов. Далее в книге излагается современное состояние наших знаний о структуре белков и нуклеиновых кислот. Главная задача этого раздела заключается в том, чтобы обобщить и упорядочить накопленную в данной области огромную информацию.

Затем излагаются данные относительно других типов биополимерных структур, таких, как полисахариды и мембраны. При этом особое внимание обращается на их отличие от белков и нуклеиновых кислот. В заключение рассматриваются взаимодействия и силы, которые приводят к образованию упорядоченных биологических структур. Знание этих сил важно для понимания и предсказания закономерностей образования трехмерных структур белков и нуклеиновых кислот.

Глава 1 Стратегия и тактика биофизической химии

1.1. Уровни структурной организации биологических макромолекул

Молекулярные массы объектов, рассматриваемых в настоящей книге, лежат в пределах от 10^3 до 10^{12} дальтон. Во многих случаях такие системы состоят из большого количества более мелких частей, все атомы которых соединены ковалентными связями. Система в целом будет называться молекулой, если имеется определенное стехиометрическое соотношение между частями, из которых она состоит, и если при физиологических условиях система достаточно устойчива и не проявляет тенденции к самопроизвольной диссоциации.

Рассматриваемые молекулы можно подразделить на семь классов в порядке увеличения их размеров и сложности (табл. 1.1). По мере увеличения размеров системы ее удобно разделять на части возрастающей величины. Как видно из табл. 1.1, размер этих частей выбирается таким образом, чтобы каждая составляла от одной сотой до одной десятой всей системы.

К первым трем классам относятся молекулы, для которых структурная информация может быть в принципе получена с атомным разрешением. Здесь имеет смысл постановка вопроса о структуре и функции таких молекул на уровне локализации отдельных атомов или звеньев (выделенных групп атомов) и их поведения.

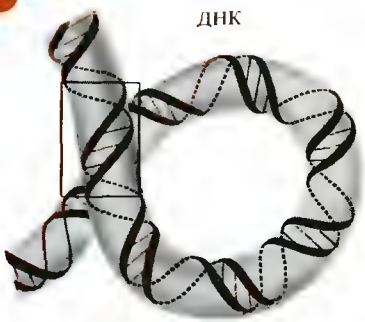
К последним четырем классам относятся столь сложные системы, что даже если бы их структура на атомном уровне была определена, такую информацию было бы чрезвычайно трудно использовать. Например, структура с мол. массой 10^6 состоит из приблизительно 10^5 атомов. Ее описание, содержащее информацию о типе и координатах каждого атома, было бы подобно описанию города путем перечисления фамилий и адресов его жителей. Удобнее подразделить систему на отдельные субъединицы или участки и затем обращаться к индивидуальным ее звеньям или атомам лишь в том случае, если их функции представляют особый интерес. Это все равно, что разделить город на районы или кварталы и затем следить за действиями лишь некоторых наиболее важных личностей, например мэра города.

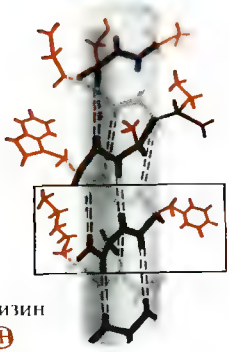
Применение при описании биополимеров таких терминов, как «субъединица», «звено», «молекула», «участок» или «группа» сопряжено с некоторой неопределенностью. Существует иерархия структурных единиц и структурных характеристик. Поскольку структурная организация белков, нуклеиновых кислот и сложных нековалентных агрегатов, таких, как наружная мембрана клетки, различна, не существует единой оптимальной номенклатуры для всех трех типов структур. Более того, даже для белков нет общепринятой

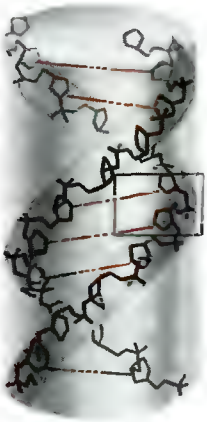
РИС. 1.1. Уровни структурной организации биологических макромолекул. Головка бактериофага или другая дискретная единица, состоящая из белков или нуклеопротеидов, отвечает четвертичной (4°) структуре. Трехмерная структура каждого отдельного компонента одного из таких комплексов является третичной (3°) структурой. Отдельные спиральные участки в такой структуре называются вторичной структурой (2°). Химическая структурная формула, отражающая лишь систему химических связей, но не пространственное расположение атомов, — это первичная (1°) структура. (Рисунок Ирвинга Гейса).

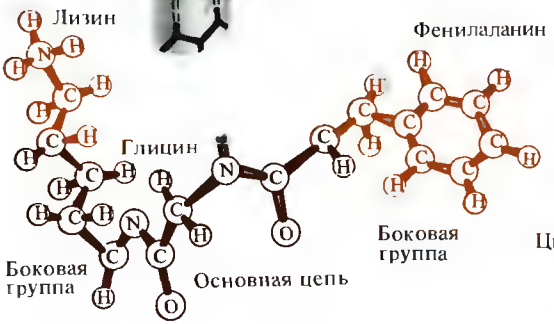
ВИРУС
ДНК
Белковая оболочка

БЕЛОК


ДНК


α - Спираль


Двойная спираль


Лизин
Фенилаланин
Глицин
Боковая группа
Боковая группа
Основная цепь

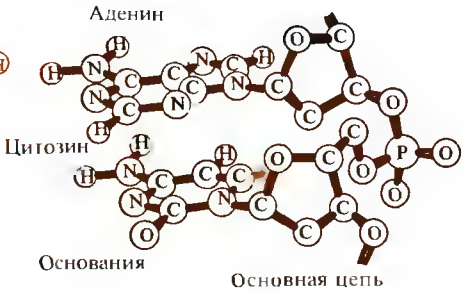

Аденин
Цитозин
Основания
Основная цепь

Таблица 1.1.

РАЗЛИЧНЫЕ БИОПОЛИМЕРЫ, РАСПОЛОЖЕННЫЕ В ПОРЯДКЕ УВЕЛИЧЕНИЯ ИХ РАЗМЕРОВ И СЛОЖНОСТИ

Класс	Пример[1]	Характерный размер	Молекулярная масса	Субъединицы, рассматриваемые при подробном описании	Число субъединиц
Олигомеры[2]	Актиномицин D	Сфера диаметром 20 Å	10^3 — 10^4	Атомы (или остатки)	10^2 (или 10)
Малые белки	Химотрипсин	Сфера диаметром 40 Å	10^4 — 10^5	Аминокислотные остатки	10^2—10^3 (или 10^3 — 10^4)
Нуклеиновые кислоты	тРНК	Палочка длиной 100 Å	10^4—10^5	Нуклеотиды (или атомы)	10^2—10^3 (или 10^3—10^4)
Большие белки	Аспартат-транскарбамоилаза	Сфера диаметром 70 Å	10^5—10^7	Субъединицы или ковалентные цепочки	10—10^2
Малые комплексы	Рибосома	Сфера диаметром 200 Å	10^5—10^7	То же	10—10^2
Большие комплексы	Мембраны, вирусы	Сфера диаметром 1000 Å	10^7 — 10^{12}	Области, фрагменты, компоненты	10—10^2
Интактная ДНК	ДНК *E. coli*	«Шланг» длиной 0,1 см	10^7—10^{12}	Участки, фрагменты, компоненты	10—10^2

[1] Все эти примеры подробно обсуждаются в разных частях книги.
[2] Приставка «олиго», часто употребляемая в этой книге, обычно относится к структурам, состоящим из 2 — 25 единиц.

терминологии. Здесь мы приведем наиболее удобный, с нашей точки зрения, способ классификации уровней макромолекулярной структуры. Следует иметь в виду, что вводимые нами определения не являются жесткими, поскольку характерные черты многих биологических структур не укладываются, строго говоря, в ту или иную категорию. Рассматриваемые уровни структурной организации схематически изображены на рис. 1.1.

Мономерными звеньями, из которых строятся белки и нуклеиновые кислоты, являются соответственно аминокислотные остатки и нуклеотиды. Их *атомная структура* известна. Это означает, что известны тип каждого атома и его усредненное положение в пространстве, а также усредненное распределение электронной плотности. Эти структуры не являются абсолютно жесткими и обладают известной подвижностью, но молекулярные колебания и вращения вокруг одиночных связей обычно не слишком сильно искажают общую картину, в особенности в упорядоченных, кристаллических структурах. Под *конфигурацией* мономериого звена (в том смысле, в каком этот термин используется в органической химии) поиимают его атомную структуру, в том числе известную стереохимию асимметричных центров. *Конформация* отвечает более полному описанию и включает сведения о предпочтительных ориентациях групп, способных перемещаться вследствие внутреннего вращения. Коиформация определяется усреднением по стерически допустимым атомным структурам и отвечает наиболее полному описанию, которое можно осуществить для молекул в растворе.

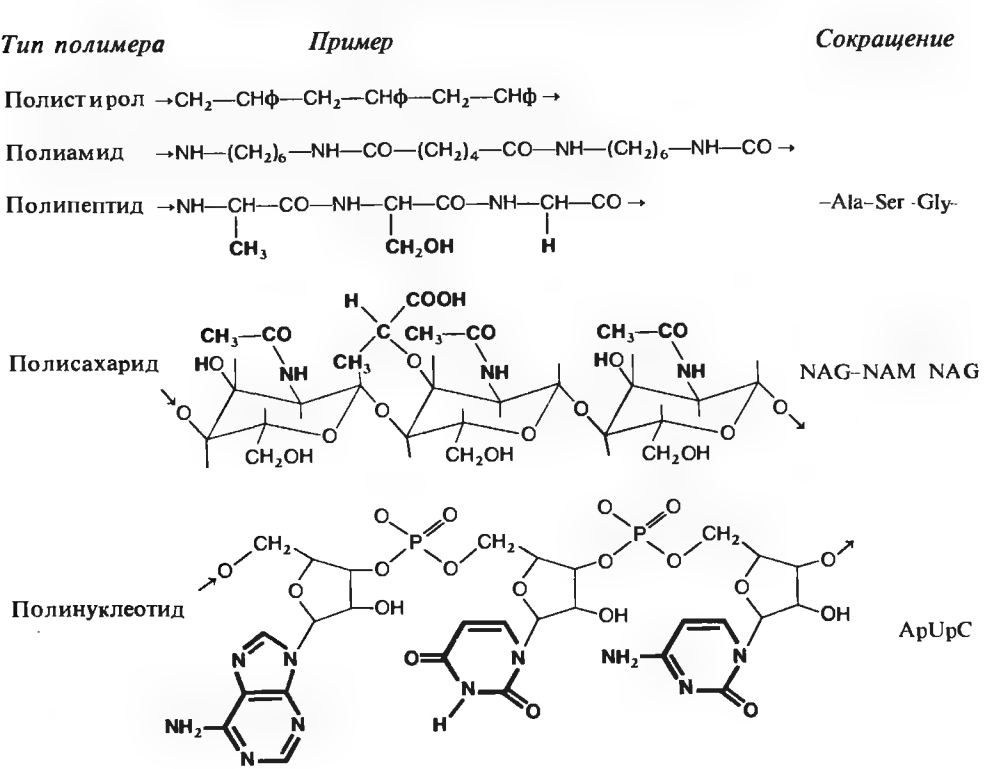


РИС. 1.2. Типичные линейные полимеры. Некоторые синтетические полимеры (например, полиамиды) не обладают определенным направлением цепи, в то время как другие синтетические и большинство биологических полимеров характеризуются расположением мономерных остатков по принципу «голова к хвосту». Основная цепь изображена обычным шрифтом, боковые группы — жирным.

ПЕРВИЧНАЯ СТРУКТУРА, Т.Е. ПОСЛЕДОВАТЕЛЬНОСТЬ ЗВЕНЬЕВ

Белки и нуклеиновые кислоты — неразветвленные линейные полимеры. Структура их ковалентной цепочки может быть записана в виде $R_1\ R_2 \ldots R_i \ldots R_n$, где R_i означает i-е звено в цепи. Особенностью белков, нуклеиновых кислот и большинства полисахаридов, отличающей их от многих синтетических полимеров, является то, что их мономерные единицы всегда соединены по принципу «голова к хвосту», т.е. цепи имеют определенное направление. Следовательно, обозначения $R_a\ R_b$ и $R_b\ R_a$ отвечают молекулам, имеющим разную химическую структуру, а концы полимерной цепи $R_1 \ldots R_n$ имеют разные химические свойства. Для описания химической структуры необходимо указать *последовательность* всех остатков. Полная ковалентная химическая структура может содержать несколько полимерных цепей, а последние могут быть соединены между собой поперечными ковалентными мостиками; подобные же мостики иногда соединяют части одной и той же цепи. Полная ковалентная структура называется *первичной структурой* (структурой 1°). Для ее описания необходимо указать конфигурацию всех асимметричных центров как основной цепи полимера, так и боковых групп каждого мономерного звена.

С помощью тонких химических методов была установлена первичная структура сотен белков и нуклеиновых кислот. Сводки первичных структур регулярно публикуются

(Dayhoff, 1975). Однако написание химической формулы для всей первичной структуры полимера требует чрезвычайно много времени, поэтому обычно применяют сокращения, используя алфавитный код. Как правило, каждый из нуклеотидов обозначают одной буквой, а каждую из аминокислот — тремя (рис. 1.2)[1]. Термины *последовательность* и *первичная структура* часто рассматриваются как взаимозаменяемые. Они действительно идентичны в случае структур, состоящих из одиночной цепи, не связанных поперечными мостиками. Следует также отметить, что ни тот, ни другой термин не подразумевает знания конформации отдельных звеньев или основной цепи.

ВТОРИЧНАЯ СТРУКТУРА ОТВЕЧАЕТ СПИРАЛЯМ, ПОСТРОЕННЫМ ИЗ МОНОМЕРНЫХ ЗВЕНЬЕВ

Многие цепи биополимеров образуют локально упорядоченные трехмерные структуры. В случае линейных полимеров, построенных по принципу «голова к хвосту» и содержащих асимметричные мономерные звенья, единственно возможным видом симметричной трехмерной упорядоченной структуры является спираль. Она имеет винтовую ось симметрии. В цилиндрических координатах с осью z, совпадающей с винтовой осью, положение j-го звена основной цепи простого спирального полимера, состоящего из $n + 1$ звеньев, определяется из уравнений

$$z_j = jz_0 + \delta_z \qquad (1.1)$$
$$x_j = r\cos(2\pi jz_0/P + \delta) \quad y_j = r\sin(2\pi jz_0/P + \delta)$$

где $j = 0,1, \ldots, n$. Как видно из рис. 1.3, r — расстояние данного звена от оси z в плоскости $x - y$; z_0 — расстояние между соседними звеньями по вертикали, вдоль оси z; P — шаг спирали — расстояние по вертикали,на котором спираль совершает один полный оборот. Постоянные δ_z и δ определяют положение нулевого звена. В более сложной спирали каждое звено состоит из нескольких атомов. P и z_0 при этом остаются одинаковыми для всех атомов. Однако r и фазы δ_z и δ хотя и являются константами для каждого звена, различны для разных атомов, входящих в данное звено.

Следует рассмотреть несколько предельных случаев. Если $r = 0$, получается простая линейная структура, порождаемая операцией трансляции $z \to z + z_0$, $x \to x$, $y \to y$. Точно такая же линейная структура получается в случае $r \neq 0$, но $P = z_0$. Когда $P = 2z_0$, координата y каждого атома каждого звена колеблется между значениями $r\sin\delta$ и $-r\sin\delta$, в то время как координата x колеблется между $r\cos\delta$ и $-r\cos\delta$. Таким образом, получается складчатая, зигзагообразная структура, напоминающая меха гармони; в этом легко убедиться, положив $\delta = 0$.

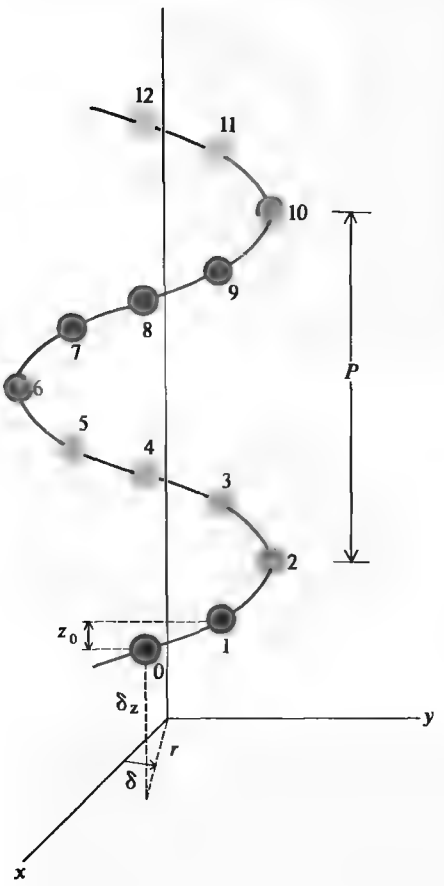
Величинна P/z_0 — число звеньев на один виток спирали. Оно не обязательно должно быть целым. Спираль с n звеньями на виток называется спиралью n-го порядка; таким образом, рассмотренная складчатая структура является спиралью 2-го порядка; полипептидная α-спираль — это спираль порядка 3,6.

Из двух или более отдельных спиральных цепей могут быть построены другие спирали. В природе часто встречаются двойные и тройные спирали. Их также можно описать уравнением (1.1), но теперь δ и δ_z для атомов разного типа в каждой цепи могут различаться. В одних многоцепочечных спиралях отдельные ковалентные цепочки имеют одинаковое направление, в других направлены навстречу друг другу. Согласно модели Уотсона и Крика, ДНК — это двойная спираль 10-го порядка с антипараллельно направленными цепями. Слоистые ß-структуры полипептидов — многоцепочечные спирали 2-го порядка, в

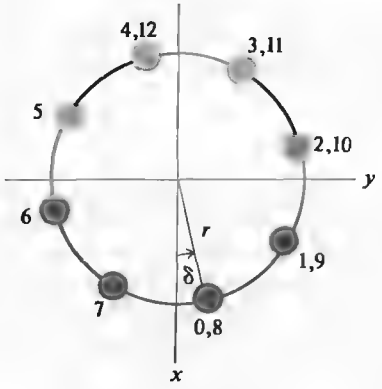
[1] По счастливому совпадению оказалось, что природа использует аналогичный способ кодирования.

которых одни цепи идут параллельно друг другу, а другие — антипараллельно. В многоцепочечных спиралях имеется симметрия, позволяющая найти координаты остатков одной цепи, если известны координаты другой. Заметим, что, поскольку белки и нуклеиновые кислоты являются сополимерами и, вообще говоря, не образуют периодической последовательности боковых групп разных типов, все операции симметрии, описанные выше, относятся только к основной цепи, но не к боковым группам.

Термин *вторичная структура* (структура 2°) относится к той части макромолекулы, которая находится в спиральном состоянии, и определяется как перечень участков первичной структуры, входящих в спираль того или иного типа. В более общем виде ее можно определить как совокупность всех трехмерных участков, имеющих упорядоченную, локально-симметричную структуру основной цепи. Для белка, например, необходимо указать, какие звенья входят в α-спирали и какие в ß-слои, а также какова полярность цепей. Участки, не имеющие форму спирали, часто называются для удобства описания «клубкообразными», хотя на самом деле по своей структуре они могут сильно отличаться от беспорядочного клубка. Аналогично в случае нуклеиновых кислот должно быть указа-



Вид сбоку



Вид сверху

РИС. 1.3. Геометрия простой спирали. Ось спирали совпадает с осью z. P — шаг спирали, z_0 — расстояние по вертикали между остатками. Положение нулевого остатка определяется величинами δ (задающей его положение в плоскости $x - y$) и δ_z (высотой над этой плоскостью). Все одинаковые повторяющиеся элементы структуры находятся на одинаковом расстоянии r от оси спирали. Эта величина называется радиусом спирали. Радиус спирали четко обозначен на виде сверху. На рисунке изображена спираль восьмого порядка. Это значит, что координаты x и y остатков, удаленных друг от друга на расстояние $8z_0$ по вертикали, совпадают.

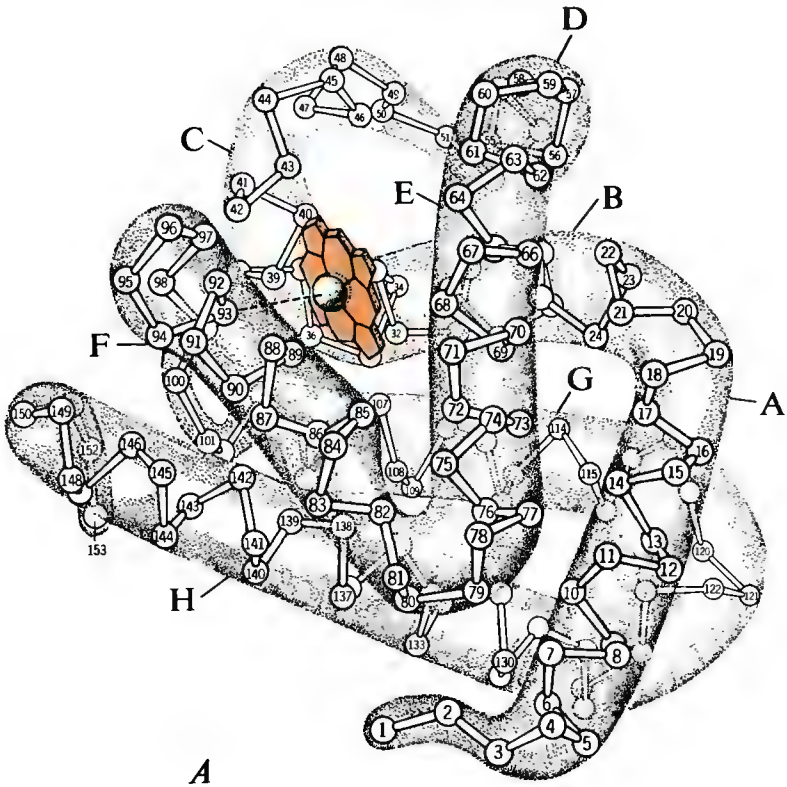
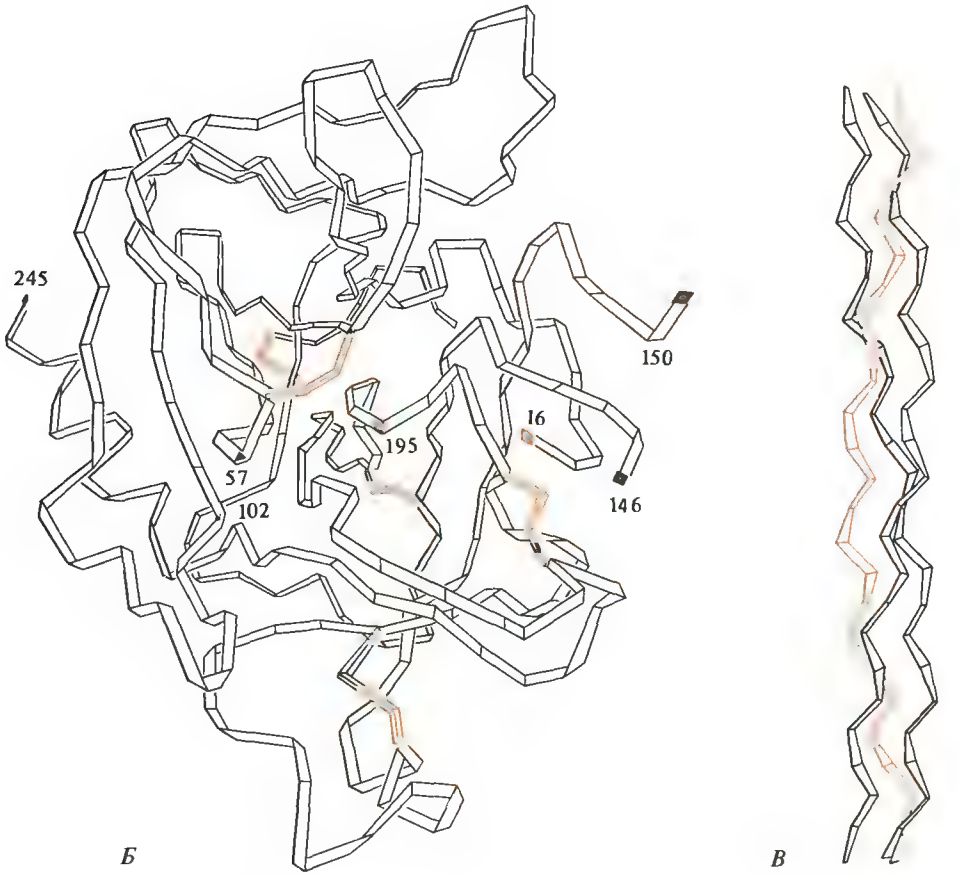


РИС. 1.4. Три типичные третичные структуры белков (показаны положения лишь α-атомов углерода). *А*. Миоглобин состоит в основном из α-спиралей, обозначенных латинскими буквами от А до Н, и содержнт группу гема (окрашена), связывающую кислород. (Рисунок Ирвинга Гейса). *Б*. α-Химотрипсин содержит три полипептидные цепи и на первый взгляд лишь один небольшой α-спиральный участок на С-конце цепочки С. Но для него характерно также иаличие значительного количества складчатых структур и четырех дисульфидных мостиков (окрашены). *В*. Часть спирали коллагена, состоящей из трех полипептидных цепей, каждая из которых находится в спиральной форме — так называемый поли-L-пролин II. За счет взаимного переплетения этих спиралей образуется сверхспираль, определяющая третичную структуру.

но для каждого данного участка последовательности, входит ли он в одиночные, двойные или тройные спирали или нет. В настоящее время известен примерно десяток дискретных типов вторичной структуры для белков и нуклеиновых кислот и еще несколько для полисахаридов. Если молекула укладывается в рамки одной совершенной вторичной структуры, например представляет собой α-спираль или двойную уотсон-криковскую спираль, то описание вторичной структуры определяет точную конформацию основной цепи, а в случае периодической последовательности, вероятно, также и боковых цепей. Для молекул, разные части которых имеют разную вторичную структуру, термин «конформация» применяется также в смысле всех имеющихся сведений об их общей структуре.



ТРЕТИЧНАЯ СТРУКТУРА, Т.Е. ПРОСТРАНСТВЕННОЕ РАСПОЛОЖЕНИЕ ЗВЕНЬЕВ

Третичная структура (структура 3°) белка или нуклеиновой кислоты — это полная трехмерная структура одной неделимой структурной единицы. Для белка такой единицей является обычно частица, все атомы которой ковалентно связаны между собой. Она может быть одним полипептидом или состоять из нескольких полипептидных цепей, связанных между собой ковалентно. Для нуклеиновой кислоты это либо одна ковалентная цепь, как в случае большинства молекул РНК, либо две комплементарные цепи, как в случае большинства молекул ДНК. Для описания третичной структуры нужно описать не только локально-симметричные структуры (2°-структуры), но также, насколько это возможно, и расположение в пространстве всех звеньев.

На рис.1.4 показаны три характерные третичные структуры белков. Обычно из таких рисунков трудно извлечь необходимую осмысленную информацию, так как они содержат слишком много деталей. Поэтому часто приходится прибегать к схематическому изображению третичной структуры, показывая только основную цепь полимера или иногда еще указывая место каждого остатка, например в случае белков — положение α-углерода. Это сделано на рис. 1.4 для того, чтобы подчеркнуть три типа различных характерных осо-

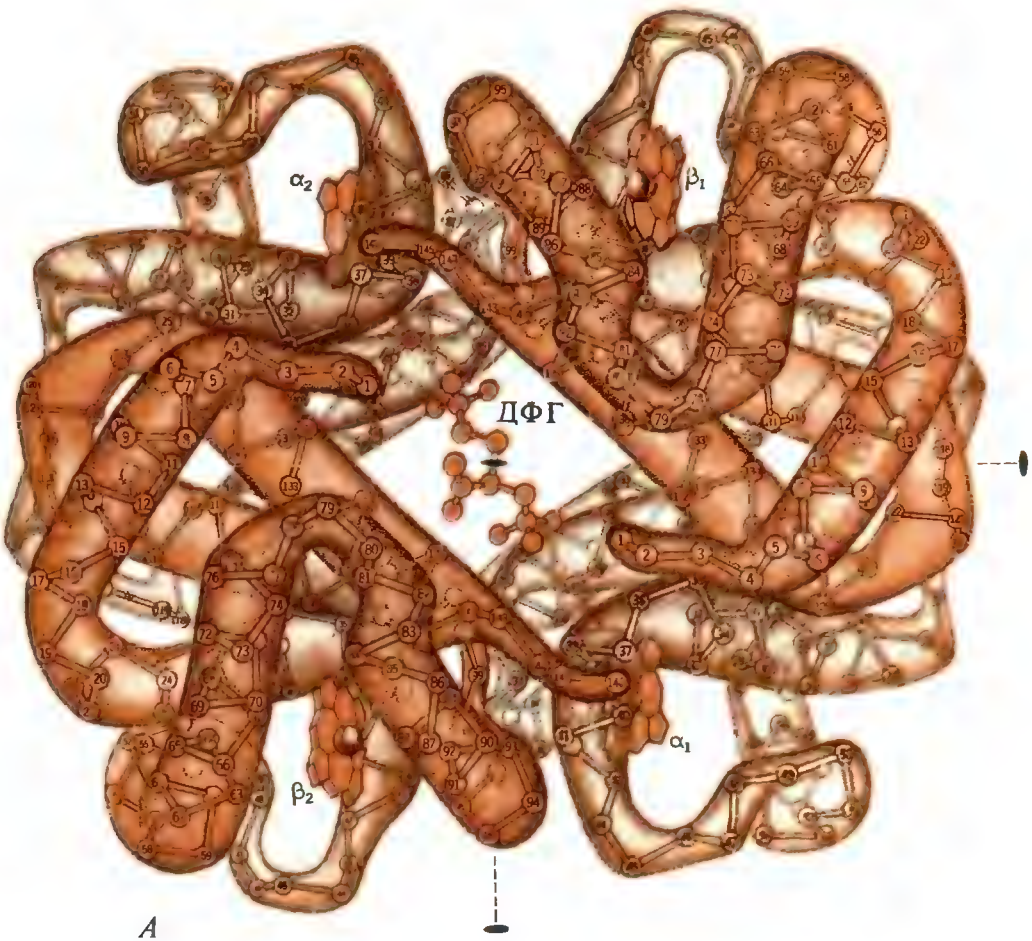
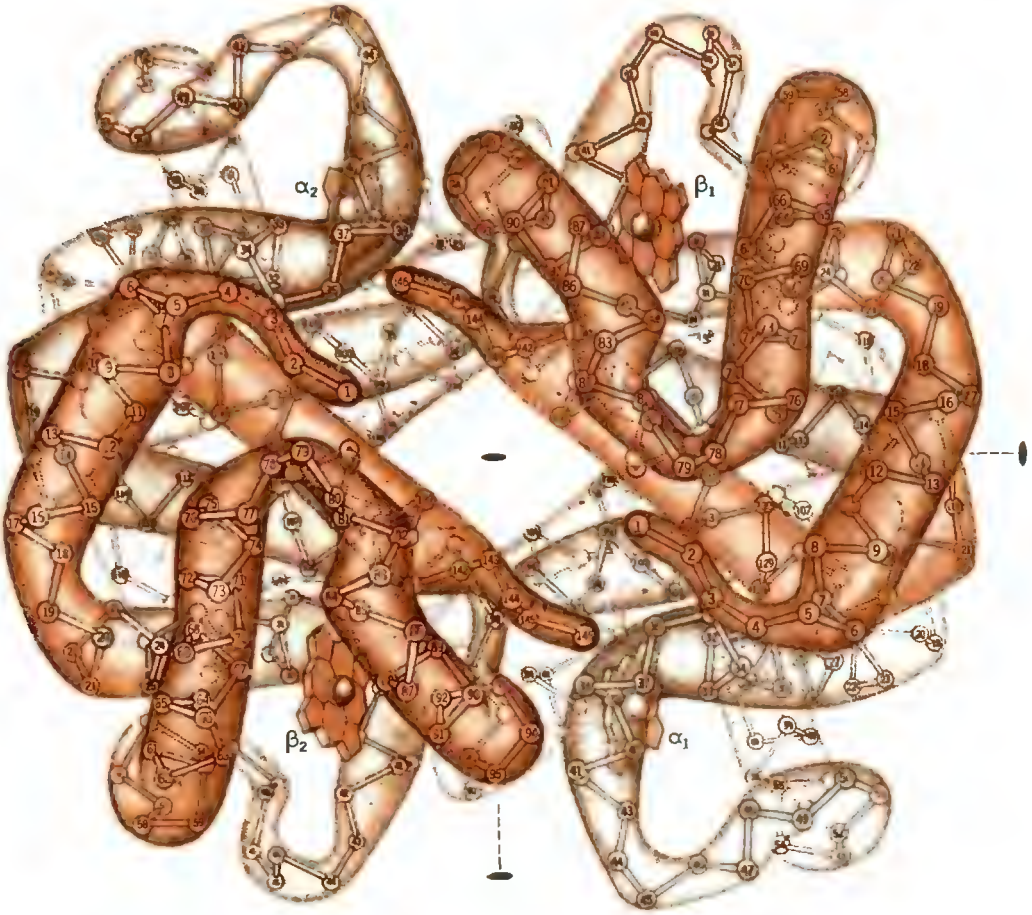


РИС. 1.5. Четвертичная структура гемоглобина. Он состоит из четырех субъединиц, с каждой из которых связан гем. Имеется ось симметрии второго порядка, расположенная в середине молекулы н перпенднкулярная плоскости рисунка. Поворот на 180° переводит α_1 в положение α_2 и β_1 — в положение β_2. Поскольку пепи α и β почти идентичны, имеются также две оси симметрии псевдовторого порядка, идущие в горизонтальном и вертикальном направлениях в плоскости рисунка; при повороте

бенностей, определяющих третичную структуру. В миоглобине почти все остатки образуют спиральную вторичную структуру, так что описание третичной структуры сводится к описанию пространственной ориентации этих спиральных участков. Химотрипсин содержит очень малую долю различимой вторичной структуры, и здесь даже грубое описание третичной структуры должно содержать информацию о положении каждого остатка. Для третичной структуры третьего типа характерна симметричная организация, отвечающая дальнему порядку. Например, благодаря периодическим нарушениям простой спиральной структуры могут образовываться сверхспирали, как одно-, так и многоцепочечные. Пример трехцепочечной сверхспирали коллагена приведен на рис.1.4.

Термин *конформация* употребляется как синоним третичной структуры. Это обусловлено тем, что если известна конформация (т.е. угол вращения вокруг всех простых связей в полимере), то тем самым известна и третичная структура, и обратно.



Б

вокруг одной из них α_2 переходит в β_1 и α_1 — в β_2, а вокруг другой — α_1 в β_1 и α_2 в β_2. *А*. Дезоксигемоглобин. Показана одна связанная молекула дифосфоглицерата (ДФГ), стабилизирующая эту форму. *Б*. Оксигемоглобин. Отметим, что субъединица повернулась, так что расстояние между β-цепями изменилось, а положительно заряженное место связывания дифосфоглицерата оказалось разрушенным.

При изучении третичной структуры не обязательно определять *полную* трехмерную структуру. Например, уже общая форма молекулы может рассматриваться как одна из характеристик третичной структуры. Сюда же относится и вопрос о том, находится ли данный остаток внутри или вне компактной глобулярной структуры белка или молекулы нуклеиновой кислоты. В настоящее время имеется несколько десятков белков и одна нуклеиновая кислота, для которых третичная структура известна столь детально, что это позволяет построить приближенную молекулярную модель. Каждый год удается расшифровать от 6 до 10 новых структур.

ЧЕТВЕРТИЧНАЯ СТРУКТУРА, Т.Е. РАСПОЛОЖЕНИЕ СУБЪЕДИНИЦ

Высшим уровнем структурной организации, которой мы здесь рассмотрим, является *четвертичная структура* (структура 4°). Она образуется путем ассоциации (за счет неко-

валентных взаимодействий) независимых единиц третичной структуры. Субъединицы четвертичной структуры могут быть как одинаковыми, так и различными, а их расположение в четвертичной структуре — как симметричным, так и несимметричным. Пример простой четвертичной структуры мы видим в случае гемоглобинов позвоночных. Молекула этих белков состоит из четырех субъединиц; она содержит по две субъединицы двух типов — α и ß. Каждая субъединица представляет собой одну полипептидную цепь, сложенную в компактную третичную глобулярную структуру. С каждой цепью связана одна группа гема. Интактный тетрамер гемоглобина $\alpha_2ß_2$ с мол. массой 66 000 легко диссоциирует на димеры:

$$\alpha_2ß_2 \rightleftarrows 2\alpha ß$$

Дальнейшая диссоциация до мономеров требует более жестких условий. Взаимное расположение субъединиц гемоглобина зависит от того, содержат ли гемы в каждой субъединице связанный кислород. Однако, как показано на рис. 1.5, каждый αß-димер всегда связан с другим осью симметрии C_2 (поворот на 180°). Кроме того, поскольку сами по себе α- и ß-субъединицы имеют весьма сходную третичную структуру, они также связаны осью псевдосимметрии.

С более сложной четвертичной структурой мы встречаемся в случае рибосомы *E. coli*. Она содержит три молекулы РНК и приблизительно 55 разных белковых цепей. Все они, кроме одной, имеются в единственном числе, один белок — в четырех экземплярах. Интактная 70S-рибосома с мол. массой $2{,}6 \cdot 10^6$ легко разделяется на две субчастицы, неравные по величине: 30S и 50S[1]. Как видно из рис. 1.6, можно осуществить дальнейшую диссоциацию каждой из этих субчастиц с образованием целого ряда промежуточных частиц. Диссоциация заканчивается получением индивидуальных компонентов. Для таких структурных образований, как рибосомы, которые на самом деле являются некоторой иерархией уровней четвертичной структуры, иногда оказывается удобным термин «пятеричная структура», который отражает более высокую степень ассоциации. Но мы будем избегать применения этого термина.

Четвертичная структура нескольких белков, состоящих из субъединиц, в настоящее время хорошо известна благодаря рентгеноструктурным исследованиям. Сложность таких белков варьирует в широких пределах — начиная со столь простых белков, как пероксиддисмутаза (состоящая из двух идентичных субъединиц), и вплоть до таких сложных молекул, как аспартат—карбамоилтрансфераза (субъединичная стехиометрия которой описывается формулой r_6c_6).

Применение электронной микроскопии позволило обнаружить и приближенно охарактеризовать целый ряд еще более сложных четвертичных структур; при этом иногда дополнительно использовался рентгеноструктурный анализ низкого разрешения. На рис. 1.7 показан для примера очень крупный сложный белок, пируватдегидрогеназный комплекс *E. coli*. Мол. масса этого комплекса составляет примерно $5 \cdot 10^6$ дальтон. Комплекс состоит из цепей трех типов — t, p и f. Суммарный субъединичный состав пируватдегидрогеназы, по-видимому, описывается формулой $t_{24}(p_2)_{12}\,(f_2)_{12}$. Другой пример — вирус табачной мозаики, состоящий из 2130 идентичных белковых субъединиц, которые образуют спираль вокруг единственной молекулы РНК длиной 6400 нуклеотидов. Имеются некоторые данные и о четвертичной структуре еще более сложных систем — о белковых ассоциатах в волокнах поперечнополосатых мышц и о белковых субструктурах в таких сравнительно сложных вирусах, как бактериофаги T2 и T4.

---

[1] В тех случаях, когда точный состав крупных структурных образований неизвестен или когда из-за сложности частиц пользоваться подобными данными неудобно, принято применять для их обозначения коэффициент седиментации в единицах Сведберга S.

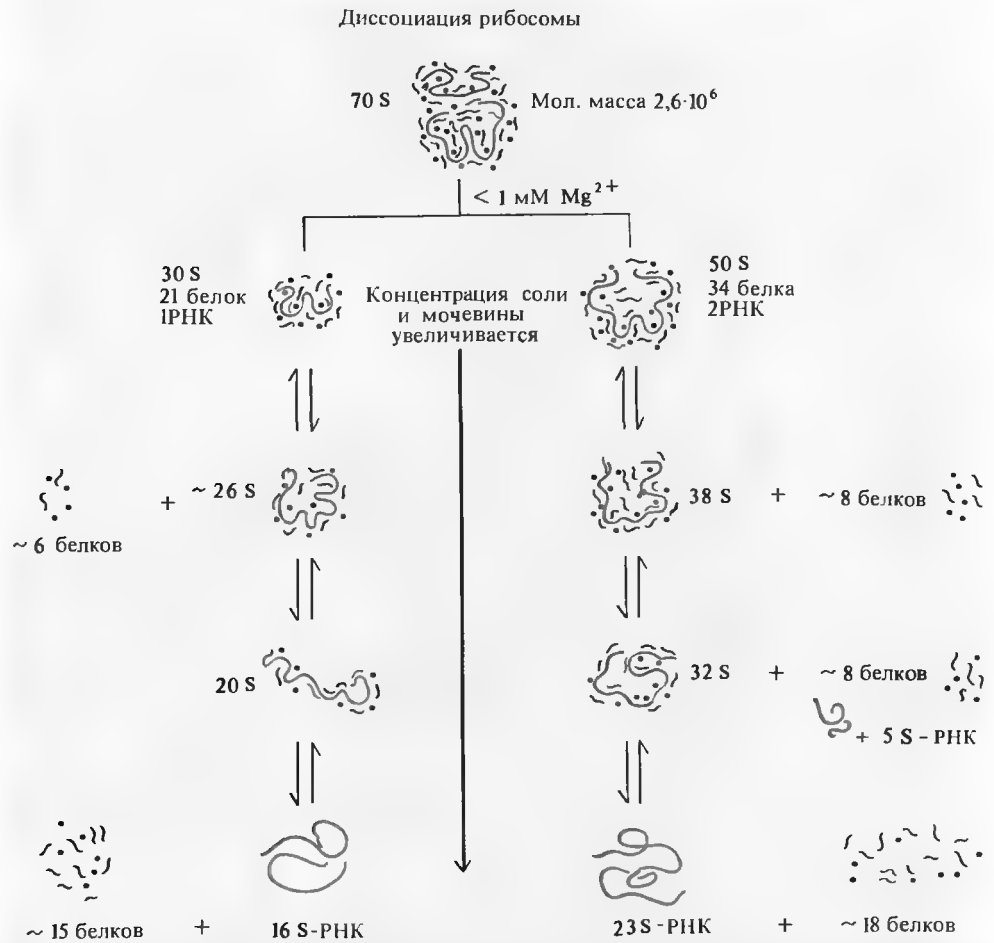


РИС. 1.6. Некоторые известные в настоящее время состояния рибосом *E.coli*, возникающие в деструктирующих условиях (по мере возрастания жесткости этих условий). Исходные 70 S-частицы содержат 55 разных белков (каждый белок, за исключением одного, представлен одной копией) и три молекулы рРНК.

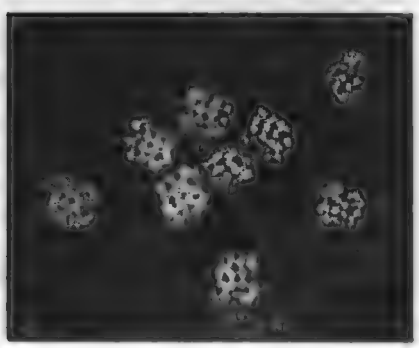


РИС. 1.7. Электронно-микроскопическая фотография пируватдегидрогеназного комплекса *E. coli*. Негативное контрастирование. (Любезно предоставлена д-ром Ридом.)

1.2. Некоторые центральные вопросы биофизической химии

После того как мы познакомились с терминологией, используемой при описании структуры больших молекул, можно сформулировать ряд вопросов, касающихся тех или иных молекул и их ассоциатов, которые часто возникают у исследователей, работающих в области биофизической химии. Не на все эти вопросы можно ответить и не все они могут быть поставлены применительно к любой системе. Задача заключается в том, чтобы определить, какие из них действительно важны для изучения той или иной системы, а затем разработать эффективную стратегию их решения.

КАЧЕСТВО ПРЕПАРАТОВ

Чтобы узнать, подходит ли данная молекула или система для детального изучения, следует ответить на следующие основные вопросы:

1. Можно ли считать препарат чистым? Является ли он гомогенным по составу? Например, содержит ли образец полимерные молекулы лишь одной молекулярной массы? На этот вопрос часто можно ответить с помощью измерения размеров молекул, используя такие методы, как ультрацентрифугирование, электрофорез и хроматографию (которые будут подробно описаны в гл. 11 и 12). Те же методы могут быть использованы и для выделения одного компонента из смеси макромолекул. При более детальном изучении гомогенности можно использовать также химический анализ, часто с применением спектроскопических измерений. Например, нередко возникают вопросы, не содержит ли препарат белка примеси нуклеиновых кислот, нет ли в нем ковалентно связанных с ним сахаров (и если есть, то сколько), состоит ли белок из одной непрерывной полипептидной цепи (и если это так, то какова ее длина). Отметим, что далеко не всегда стоит задавать вопрос об отсутствии загрязнения белка или нуклеиновой кислоты малыми молекулами, так как последние неизбежно присутствуют в препаратах. Большинство биополимеров — полиэлектролиты и поэтому находятся в окружении противоионов. Иногда бывает важно тщательно проконтролировать тип противоионов, которые присутствуют в образце. Очень часто для нормального функционирования макромолекулы бывает необходимо присутствие или отсутствие тех или иных малых молекул.

2. Является ли препарат нативным? Нет ли в нем повреждений и не потеряна ли важная его часть при выделении? Важно знать, обладает ли выделенная молекула in vitro той же структурой и теми же свойствами, что и в живой клетке, где она была синтезирована и где большинство биополимеров обычно функционирует. Мы знаем, что во многих случаях ответ будет безусловно отрицательным. In vivo многие молекулы упакованы в сложные структуры. Последние настолько сложны, что их исследование физическими методами практически невозможно. Кроме того, их целостность так легко нарушить, что большинство методов выделения молекул путем разрушения клетки разрушает и эти структуры. Например, ДНК in vivo, по-видимому, никогда не бывает свободной, подобно выделенной молекуле в растворе. Она всегда взаимодействует с белками и, очевидно, упакована в организованные или упорядоченные структуры в клетках как прокариот, так и эукариот. Даже в простейших вирусах молекула ДНК как-то уложена. ДНК внутри бактериофага Т2 плотно упакована и занимает бóльшую часть внутреннего объема головки фага. Как показано на рис. 1.8, осмотический лизис, производимый для выделения ДНК из фага, приводит к огромному увеличению общего объема, занимаемого молекулой ДНК, хотя истинный (собственный) молекулярный объем ДНК остается, по-видимому, прежним.

Поскольку условия in vivo невозможно полностью воспроизвести in vitro, приходится прибегать к одному из нескольких вариантов проведения исследований. Иногда можно изучать отдельные классы молекул внутри клетки. Для этого требуется, чтобы изучаемые

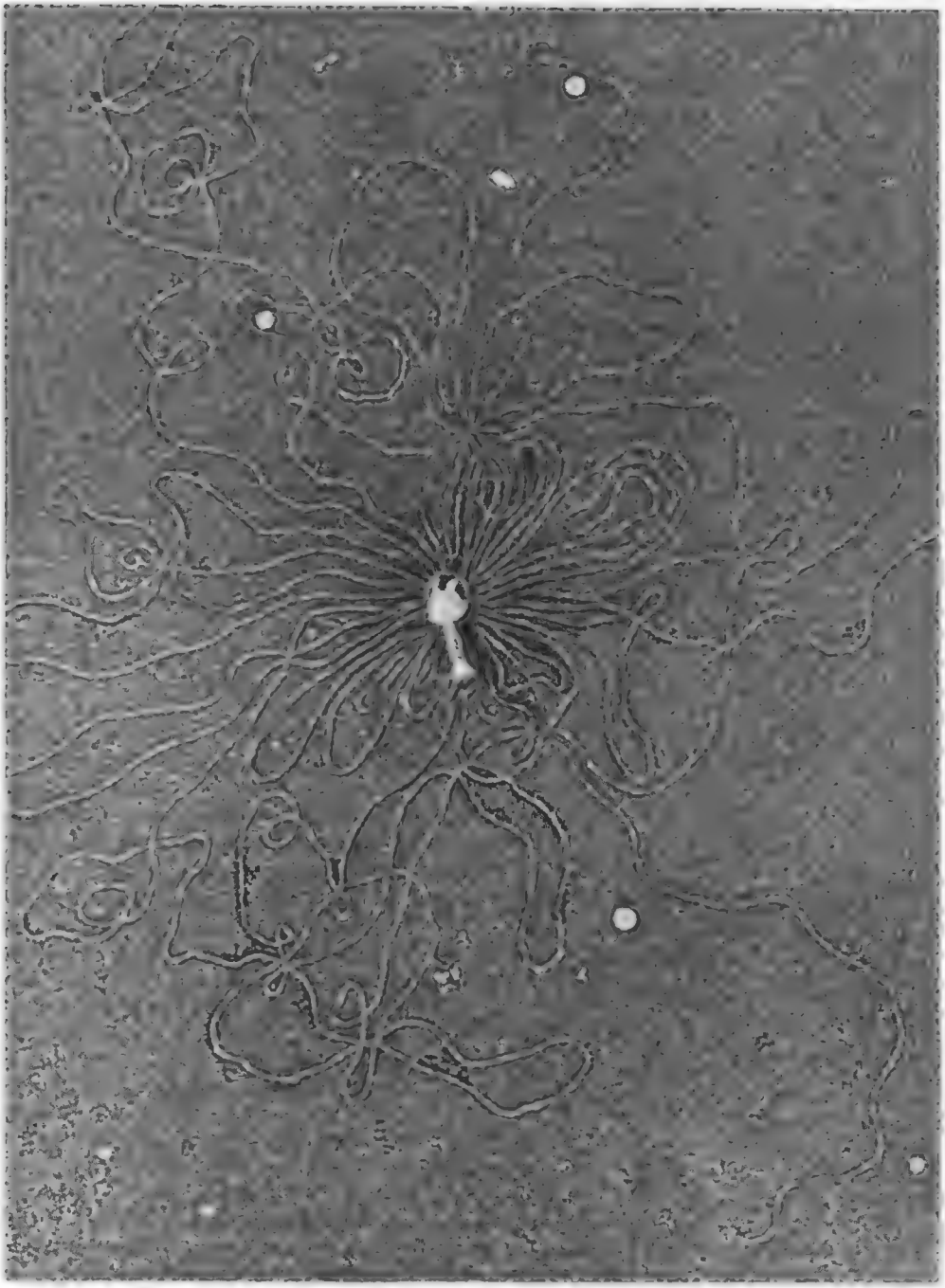


РИС. 1.8. Бактериофаг Т2, подвергнутый осмотическому шоку для освобождения ДНК. До сих пор неизвестно, каким образом огромная молекула ДНК была исходно плотно свернута для упаковки внутри головки фага. [A.K.Kleinschmidt, D. Lang, D. Jacherts, R.K. Zahn. Biochim. Biophys. Acta, 61, 857 (1962).]

молекулы либо составляли большую долю всей массы клетки (как, например, гемоглобин в эритроцитах), либо обладали особыми физическими или химическими свойствами, позволяющими наблюдать их на фоне всех остальных компонентов клетки (как, например, в случае некоторых парамагнитных или интенсивно окрашенных молекул). С другой стороны, иногда удается увеличить количество нужных молекул внутри клетки с помощью химических и биологических воздействий или ввести своего рода зонд или метку, позволяющую проводить прямые физические наблюдения. Эти методы очень важны, но их разработка находится пока в начальной стадии.

Если молекулу невозможно исследовать in vivo, то необходимо постоянно отдавать себе отчет в том, что in vitro ее свойства могут быть иными. Различия могут быть такими разительными, как в случае фаговой ДНК, рассмотренном выше, или очень небольшими, но не менее важными. В течение нескольких десятилетий гемоглобин изучали гораздо более интенсивно, чем любой другой биополимер. Об этом белке известно очень многое, в частности установлена его полная трехмерная структура, но его функциональное поведение (связывание кислорода) in vitro было по неизвестным причинам иным, чем in vivo. Более того, оно различалось в разных лабораториях. В конце концов было обнаружено, что сродство гемоглобина к кислороду резко изменяется под влиянием малых молекул дифосфоглицерата, присутствующих в эритроцитах. В разных лабораториях при очистке гемоглобина теряли различное количество дифосфоглицерата.

Серьезной проблемой является и повреждение макромолекул при их очистке. Большинство клеток содержит ферменты, способные гидролизовать или как-либо иначе модифицировать белки и нуклеиновые кислоты. Обычно эти ферменты изолированы или их активность подавлена иным способом, так что они не способны атаковать компоненты клетки, но они могут освобождаться при разрушении клетки.

3. Получаются ли при использовании препарата повторяющиеся результаты? Обнаруживают ли молекулы идентичные свойства при работе с разными препаратами? Изменения в состоянии живых организмов, из которых молекулы были выделены, или неизбежные небольшие вариации при работе с особо чувствительными системами могут приводить к получению измененных структур. Хорошим контролем постоянства препаратов являются функциональные тесты. Имеется и ряд щадящих физических измерений, достаточно чувствительных для того, чтобы служить контролем качества препаратов. Если известна причина неидентичности препаратов, иногда ее удается устранить. Например, повреждение РНК за счет ферментативного расщепления может быть сведено к минимуму путем выделения ее из мутантных бактериальных штаммов, обладающих дефектными или неактивными расщепляющими РНК ферментами; присущая живым организмам вариабельность может быть резко уменьшена за счет использования животных клеток, растущих в культуре в стандартных условиях. Генетические вариации могут быть устранены путем использования клонированных клеток, происходящих от одного генетически чистого предшественника.

ВОПРОСЫ О СТРУКТУРЕ

После того как мы получили чистую и воспроизводимую систему, можно поставить множество интересных вопросов, относящихся к ее молекулярным характеристикам.

1. Какова ее структура? Это естественный вопрос, и на него можно получить ответ на разных уровнях, в зависимости от природы самой системы, от затраченных усилий и от количества имеющегося материала. Определение первичной структуры — это в первую очередь биохимическая, а не физическая задача. Однако знание первичной структуры является обязательным предварительным условием многих физических исследований (см. введение к тому II). Иногда физические методы оказываются полезными при определении

последовательности мономерных звеньев. Несколько ниже будут приведены соответствующие примеры.

Информацию о вторичной структуре легче всего получить из спектроскопических исследований. Оптические методы, обсуждаемые в гл. 7 и 8, могут дать надежные сведения о характере вторичной структуры белка и нуклеиновой кислоты и о количестве различных форм. С их помощью, вообще говоря, нельзя определить, какие именно мономерные звенья входят в те или иные участки вторичной структуры. Детальную информацию об этом для нуклеиновых кислот может дать ЯМР в сочетании с другими методами (гл.9).

Полная третичная структура биологической макромолекулы может быть установлена в настоящее время лишь с помощью рентгеновской кристаллографии и некоторых тесно связанных с ней дифракционных методов (гл. 13 и 14). Все эти методы требуют включения молекул в хорошо упорядоченные кристаллические структуры, а это можно сделать лишь в случае некоторых биополимеров. Дело в том, что определенная часть биологических молекул или систем неупорядочена по самой своей природе. Такие системы невозможно исследовать с высоким разрешением с помощью дифракционных методов. Если получить достаточно совершенные кристаллы не удается, можно попытаться исследовать третичную структуру биополимера, используя сочетание ряда менее информативных методов. Электронная спектроскопия и гидродинамические методы (гл. 10-12) дают сведения о размере и форме молекулы. Ряд физических и химических подходов может дать информацию о том, доступны ли определенные звенья для взаимодействия с молекулами раствора. Если такое взаимодействие имеет место, мы можем заключить, что эти звенья расположены преимущественно на поверхности изучаемой структуры, а в противоположном случае — в глубине. Некоторые спектроскопические методы позволяют получить более детальные сведения о третичной структуре. Так, затратив значительные усилия, можно измерить с их помощью расстояние между определенными точками внутри изучаемой структуры. Трудность заключается в том, что при этом удается определить в одном опыте лишь одно из расстояний.

Другие спектроскопические методы при условии, что молекулы могут быть упорядочены, позволяют установить ориентацию определенных участков вторичной структуры или отдельных звеньев в рамках третичной структуры. В случае четвертичной структуры имеется много различных подходов, которые позволяют определить число и типы субъединиц. Обычно они сводятся к аккуратному количественному исследованию интактной системы и измерению молекулярной массы всех субъединиц. Если система достаточно велика по размерам, то геометрию расположения субъединиц можно установить с помощью электронного микроскопа. В противном случае сведения о ней могут быть получены с помощью химических методов (например, пришивание) или некоторых физических методов, дающих информацию о расстояниях. Если система состоит из малого числа субъединиц, иногда можно получить данные о структуре путем тщательного анализа гидродинамических данных.

2. Можно ли предсказать структуру? Дать утвердительный ответ на этот вопрос было одной из главных целей биофизической химии в течение многих лет. Рассмотрим прежде всего вторичную структуру. Говоря о предсказании, необходимо рассмотреть два аспекта. Можно ли предсказать или количественно объяснить, исходя из заданной химической структуры основной цепи полимера, наиболее устойчивую геометрию, используя для этого известные сведения о ковалентных и нековалентных взаимодействиях? Как будет показано в гл. 5 и 6, ответ на этот вопрос в основном положительный. Отдельный вопрос заключается в том, можно ли предсказать характеристики вторичной структуры, которая сформируется при заданных внешних условиях, исходя из данной первичной структуры и из определенного набора вторичных структур. Этот вопрос сводится для белка к вопросу о том, какие остатки окажутся в спиралях или слоях, а для нуклеиновой кислоты — какие остатки окажутся в двойной спирали. Для таких предсказаний необходимо знать, как каж-

дая боковая цепь изменяет энергию вторичной структуры того или иного типа. Должен иметься также способ поиска оптимальной структуры среди всех возможных.

Для нуклеиновых кислот решение этой задачи проще, чем для белков. В этом вопросе имеются значительные успехи. Основные преобладающие взаимодействия связаны здесь с образованием уотсон-криковских пар. Возможность такого спаривания непосредственно видна из рассмотрения первичной структуры. Чтобы рассчитать энергию спаривания при той или иной последовательности оснований, нужно учесть параметры менее 40 термодинамических взаимодействий. Этот вопрос будет рассмотрен в гл.23. В случае белков задача оказывается значительно более трудной. Аминокислотная последовательность сама по себе не дает таких простых следствий, как правило комплементарности при спаривании оснований. Число параметров термодинамических взаимодействий, которые должны учитываться, оказывается значительно больше. Это обусловлено тем, что число разных аминокислотных остатков равно 20, в то время как число разных нуклеотидов — всего 4. Несмотря на эти трудности, существуют, как мы покажем, более или менее надежные способы предсказания вторичной структуры белков.

На втором этапе предсказания структуры делается попытка перехода от первичной и вторичной структуры к третичной структуре. Математическая сложность этой задачи, как будет видно из дальнейших глав, огромна. Для строгого решения оказываются недостаточными возможности ЭВМ — как существующих, так и тех, которые появятся в будущем. Однако за последние несколько лет был достигнут значительный прогресс в развитии методов предсказания того, как молекула белка или нуклеиновой кислоты должна складываться в трехмерную структуру. Это удалось сделать благодаря правильно выбранным приближениям и упрощениям. Тем не менее действительно успешное предсказание третичной структуры полимера до сих пор не было осуществлено. Аналогичной, но, по-видимому, значительно более простой задачей является предсказание четвертичной структуры некоего комплекса, исходя из заданной третичной структуры его компонентов. Если бы последние были жесткими, задача могла бы быть решена путем минимизации полной энергии с помощью перебора всех разумных структур методом проб и ошибок. Такие попытки делались с целью исследования возможной структуры волокон, образующихся при агрегации гемоглобина больных серповидноклеточной анемией. Уже небольшое количество данных о структуре этих волокон существенно облегчает вычисления за счет ограничения числа геометрических переменных. Вообще говоря, в будущем методам предсказания структуры будет, очевидно, принадлежать огромная роль — может быть, не при расчетах a priori, а как способу использования всей имеющейся структурной информации с целью получения возможно более точной и детальной структуры.

Иногда удается экспериментально исследовать вопрос о степени взаимодействия различных частей структуры между собой, например установить, может ли один из спиральных участков нативного глобулярного белка выйти из этой структуры, не влияя при этом на остальные. Если на самом деле удалить этот участок из молекулы, скажем, с помощью ферментативного отщепления, останется ли он спиральным в изолированном состоянии? Как скажется это удаление на остальной структуре? Допустим, что белок или нуклеиновая кислота подвергается действию внешних факторов, например повышению температуры. Явление, состоящее в том, что малое возмущение вызывает большое изменение в третичной и вторичной структуре, называется кооперативным структурным переходом. Этот вопрос рассматривается в гл. 20 и 21. Такое превращение предполагает, что различные и (возможно) удаленные друг от друга участки молекулы могут в значительной степени взаимодействовать. Следовательно, предсказать структуру a priori будет весьма трудно из-за того, что может оказаться невозможным разбить систему на части в целях проведения расчетов.

3. Является структура жесткой или гибкой? Все молекулы в той или иной степени обладают гибкостью. В случае больших молекул нас интересует, обладают ли подвижнос-

тью боковые цепи или более протяженные участки. Некоторые макромолекулы, например гемоглобин, могут находиться в нескольких дискретных конформациях, как показано на рис. 1.5. Подобное обстоятельство часто используется для регуляции их биологических функций. Другие молекулы, например типичный денатурированный белок или беспорядочный полимерный клубок, характеризуются целым спектром конформаций, и тогда удается описать лишь усредненную структуру системы. Чтобы молекула распознавала лишь весьма специфическую мишень и связывалась с ней, в простейшем варианте она должна представлять собой жесткую матрицу, точный отпечаток мишени. Однако в ряде случаев возникновению специфических взаимодействий способствует именно гибкость молекул.

Некоторые молекулы, например многие иммуноглобулины, представляют собой набор относительно жестких образований, соединенных между собой несколькими гибкими связями. По-видимому, гибкость, свойственная таким структурам, связана со специфической функциональной ролью молекулы. Каждая молекула иммуноглобулина G(JgG) имеет два центра связывания со специфическим антигеном (рис. 1.9). Если бы иммуноглобулин G был жесткой молекулой, он мог бы связывать два антигена, расположенных на поверхности или в объеме, лишь в том случае, когда их расположение точно отвечает геометрии этих двух центров. Если свободная энергия связывания в расчете на один центр равна ΔG^0, то кажущаяся макроскопическая константа связывания IgG с одним антигеном равна $2\exp(-\Delta G^0/RT)$, где наличие множителя 2 обусловлено тем, что каждая ветвь IgG может специфически связаться с одним из антигенов. Однако, поскольку IgG обладает гибкостью, гораздо более вероятно, что после того как один из центров будет занят, второй тоже окажется занятым. Это — проявление «хелатного эффекта». Хотя энтальпия связывания со вторым центром такая же, как и с первым, потеря энтропии гораздо меньше, так как второй центр уже зафиксирован вблизи антигена благодаря тому, что первый занят антигеном. А жесткий двухвалентный иммуноглобулин способен использовать хелатный эффект для увеличения сродства к антигену лишь при том условии, что его структура случайно допускает одновременное заполнение обоих центров связывания.

Один из способов исследования области допустимых конформаций макромолекулы заключается в варьировании условий среды. Нековалентные силы, стабилизирующие макромолекулярные структуры, обсуждаются в гл. 5. Исходя из свойств этих сил, можно

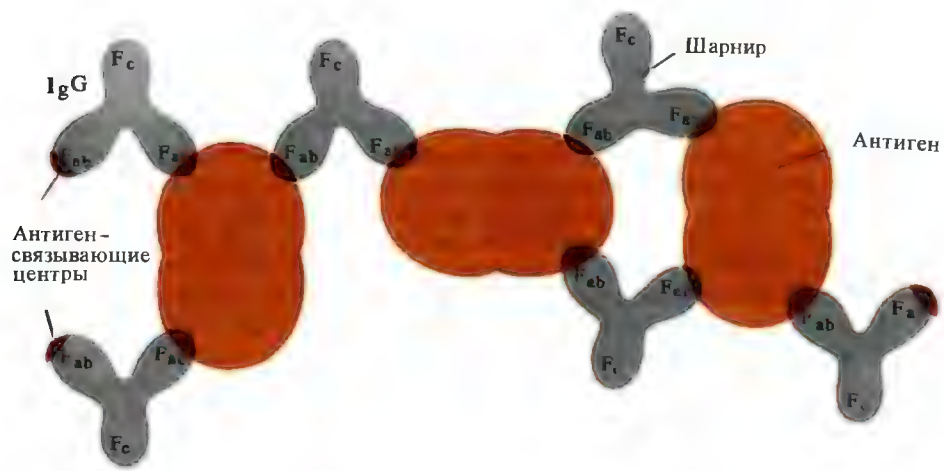


РИС. 1.9. Схема взаимодействия антител с антигенами. Несколько молекул иммуноглобулина G в типичном преципитате антиген — антитело.

заключить, что изменения температуры, pH, концентрации соли или состава растворителя способны вызвать изменения в структуре [1]. Исследование структуры в зависимости от этих термодинамических переменных часто позволяет выявить целый ряд дискретных структурных форм и измерить разность энтальпий и энтропий между этими формами. Данный вопрос обсуждается в гл. 20-23.

В принципе чрезвычайно интересно разобраться в механизме влияния изменений окружающих условий на вторичную, третичную и четвертичную структуру. Преимущественная чувствительность к определенной переменной часто дает ключ к пониманию того, какой тип нековалентных взаимодействий отвечает за преимущественное поддержание одной определенной структуры. Более специальный случай внешних возмущений — это связь между конформационными изменениями на разных структурных уровнях. Например, вызывают ли изменения в четвертичной структуре, такие, как удаление одной субъединицы, изменения в третичной структуре какой-либо другой субъединицы? И наоборот, как влияет изменение третичной структуры субъединицы на характер и стабильность той четвертичной структуры, в которую она входит?

4. Имеется ли в исследуемой структуре что-либо знакомое или она совершенно необычна? Из сказанного выше следует, что каждому структурному уровню белков или нуклеиновых кислот, по-видимому, присущи общие характерные особенности. Некоторые из этих характеристик являются результатом действия сил, определяющих структуру. Даже если мы не можем точно предсказать структуру, иногда оказывается возможным распознать некоторые ее элементы, противоречащие нашим сегодняшним представлениям. Такие структурные особенности часто оказываются очень важны для понимания функции. Например, в случае глобулярных белков можно в первом приближении предсказать особенности распределения аминокислотных остатков между поверхностью молекулы и внутренней областью. Заряженные группы, например лизин или аргинин, обычно сольватированы и окружены противоионами. Перенос изолированной боковой цепи такого типа внутрь глобулы не только приведет к потере энергии сольватации, но и будет весьма дорого стоить из-за электростатических взаимодействий. Например, энергия притяжения двух противоположно заряженных ионов примерно равна $e^2/\varepsilon r$, где e — элементарный заряд, r — расстояние между ионами, рассматриваемыми как точечные заряды, а ε — диэлектрическая проницаемость среды между ними. Вне белковой глобулы ε равно примерно 80, а внутри — приблизительно в 20 раз меньше. Поэтому, если необходимо перенести заряд внутрь, энергетически чрезвычайно выгодно поместить возможно ближе к нему еще один заряд противоположного знака. Обнаружение зарядов внутри глобулы белка подобно обнаружению оазисов в пустыне. Вероятно, они находятся там не случайно.

Сравнивая между собой те белки или нуклеиновые кислоты, структура которых изучена достаточно хорошо, можно обнаружить замечательную тенденцию: некоторые молекулы, несмотря на их различие, обладают поразительно сходной структурой, причем даже на разных уровнях. Поэтому всегда имеет смысл сравнивать вновь найденную структуру с известными структурами. Если она попадает в определенное семейство структур, подобное, например, семейству сериновых эстераз, то информация о членах этого семейства, полученная ранее, существенно ускоряет изучение новой структуры. Наблюдаемые структурные различия являются при этом решающим обстоятельством для понимания источника любых функциональных различий.

Происхождение самих этих структурных семейств, а также их распространенность остаются до сих пор неясными. Конечно, главной причиной является дивергентная эво-

---

[1] Давление и электрическое поле также важны, но изучаются реже, так как их труднее изменять в ходе эксперимента.

люция. Примитивные организмы должны были содержать значительно меньше различных белков или нуклеиновых кислот, чем современные. Простейшим ответом на всякое новое давление отбора была бы дупликация генов с последующей небольшой модификацией одного из дочерних генов, приводящей к новой функции. Действительно, первичная структура или более высокие уровни структурной организации могут использоваться для того, чтобы проследить пути эволюции организмов, подобно тому как для этого используются строение костей или другие анатомические особенности. Другое, более спекулятивное объяснение некоторых структурных гомологий, наблюдаемых у биополимеров, связано с конвергентной эволюцией. Считают, что некоторые биологические структуры являются столь оптимальными для решения данной функциональной задачи, что молекулы, исходно удаленные от этого оптимума, будут со временем приближаться к нему в ответ на давление отбора. Трудно доказать, что конвергентная эволюция внесла важный вклад в структурную картину, наблюдаемую сегодня, но также трудно полностью исключить такую возможность.

Иногда структура оказывается столь необычной, что позволяет сделать непосредственные выводы о функции. Например, под электронным микроскопом частицы белка ферритина выглядят как белковая оболочка, окружающая электроноплотное ядро. Химический анализ показывает, что оно состоит преимущественно из Fe_2O_3. Становится ясно, что ферритин является белком — накопителем железа. Оксиды железа нерастворимы в воде, и для организма было бы затруднительно транспортировать их и вообще обращаться с ними. Ферритин изящно разрешает эту проблему путем упаковки железа внутри белка.

5. Не изменены ли свойства индивидуальных звеньев? Вообще говоря, нам хорошо известны ионные равновесия, химические и физические свойства всех обычных боковых цепей, входящих в белки и нуклеиновые кислоты. Однако эти свойства часто изменяются, когда остатки оказываются включенными в упорядоченную вторичную структуру или структуры высших порядков. Это изменение может явиться источником информации о локальной структуре вблизи данного остатка. Изменение свойств можно исследовать при помощи некоторых физических методов с гораздо более высоким разрешением, чем это удается при исследовании структуры как целого методом рентгеновской кристаллографии. Например, определенные спектроскопические методы обладают высокой чувствительностью к характеру химической связи или к распределению электронов вокруг данного атома. Если этот атом чем-то выделен, он может служить зондом, поскольку его можно изучать на фоне всей макромолекулы.

Иногда ключом к пониманию функции данного звена являются аномальные химические свойства. Гидроксильные группы (например, в серине) обычно обладают лишь небольшой реакционной способностью по отношению к электрофильным реагентам. Положение изменяется только при высоких pH, когда они депротонируются. Однако давно известно, что определенный сериновый остаток в некоторых белках обладает высокой реакционной способностью даже при нейтральных pH. Это используется сериновыми эстеразами, в которых активированный серин играет роль нуклеофильного агента в их каталитическом центре. Теперь известно, что соседи упомянутого серина в третичной структуре взаимодействуют с ним таким образом, что это приводит к его сильной активации. Сначала считали, что активация связана с системой переноса заряда, стабилизирующей ионизованный серин (рис. 1.10.). Однако наиболее точные рентгеноструктурные данные, имеющиеся в нашем распоряжении, свидетельствуют о том, что наблюдаемая химическая активность серина является результатом его выгодного расположения рядом с электрофильным карбонилом связанного субстрата, а не результатом его повышенной нуклеофильности.

6. Каким образом возникает данная структура? Одним из фундаментальных разделов биологии является морфология, рассматривающая развитие и поддержание формы. В случае биополимеров механизмы, ответственные за формирование структуры, могут

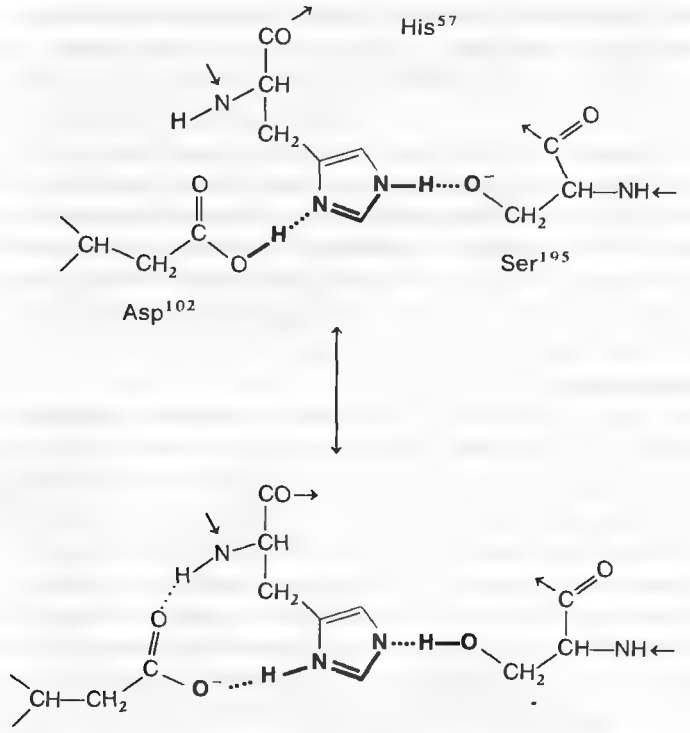


РИС. 1.10. Модель образования водородных связей между тремя остатками в активиом центре химотрипсина и других сериновых эстераз. Показаны две таутомериые структуры. Они образуют систему, сообщающую гидроксилу серина небольшой отрицательный заряд. Одиако новые данные ставят под сомнение идею о наличии водородной связи между Ser^{195} и His^{57}. [См. D. A. Matthews et al., J. Biol. Chem., 252, 8875 (1977).]

быть различными для разных структурных уровней. Что касается первичной структуры, то, за исключением образования поперечных связей, это в основном биохимическая проблема[1]. Механизм синтеза специфических аминокислотных и нуклеотидных последовательностей известен достаточно детально и здесь обсуждаться не будет. Вторичные структуры, в особенности те, образование которых связано с взаимодействием лишь между соседними (в рамках первичной структуры) звеньями, в благоприятиых окружающих условиях образуются быстро и самопроизвольно. О детальных стадиях этого процесса известно достаточно много, и этот вопрос применительно к белкам и нуклеиновым кислотам будет обсуждаться в гл.20-23.

Механизмы формирования третичной и четвертичной структур более сложны и противоречивы. Ясны лишь некоторые их аспекты. Ряд третичных структур белков и нуклеиновых кислот может in vitro образовываться самопроизвольно. Эти эксперименты проводятся по следующей схеме: биополимер с нативной третичной структурой подвергают максимально возможной денатурации, затем денатурирующее воздействие устраняют и наблюдают восстановление структуры или функциональной активности. Насколько мож-

[1] Хотя механизм недавно открытого сплайсинга РНК может быть связан и с более высокими уровнями структурной оргаиизации.

но судить, при этом образуются молекулы со структурой, идентичной во всех отношениях исходной нативной структуре. В некоторых случаях оказывается возможным наблюдать отдельные стадии ренатурации, что позволяет судить о деталях процесса образования третичной структуры. Однако, как правило, трудно сказать, чем определяется этот процесс — термодинамикой или кинетикой. Если он определяется термодинамикой, то конечная структура является термодинамически наиболее устойчивой и молекула «перебирает» весь набор этих структур. Если процесс определяется кинетикой, это означает, что перебирается лишь ограниченный набор промежуточных структур. При этом молекула в процессе достижения конечной структуры проходит через состояния, разделенные низкими кинетическими барьерами. В результате конечная структура может оказаться метастабильной и не соответствовать абсолютному термодинамическому минимуму.

Точно такое же рассмотрение и такая же неопределенность относятся к процессу образования четвертичной структуры. Ярким свидетельством экспериментального искусства многих талантливых исследователей является тот факт, что удалось осуществить in vitro самосборку ряда наиболее сложных четвертичных структур — рибосом, нуклеосом и даже целых вирусов из отдельных компонентов. Как и в случае третичной структуры, вопрос величайшего биологического значения заключается в том, проходит ли сборка in vivo тем же путем, что и in vitro. Ответить на него утвердительно очень трудно. Наиболее доказательный подход заключался бы в том, чтобы найти мутантов, дефектных по сборке, и сравнить промежуточные состояния, накапливающиеся in vivo, со стабильными промежуточными формами, которые можно получить in vitro. Однако у этих мутантов по самой их природе процесс сборки in vivo изменен, и наблюдаемые промежуточные формы могут не соответствовать тем подлинным формам, которые образуются в норме.

В некоторых случаях можно показать, что сборка in vivo должна проходить иначе, чем процесс, протекающий после простого смешивания компонентов. Следует задать два важнейших вопроса. Все ли молекулы, участвующие в процессе сборки, представлены в конечной структуре? Имеют ли все молекулы, входящие в конечную структуру, ту же первичную структуру, которую они имели в момент их присоединения к комплексу in vivo? Если хотя бы на один вопрос будет дан отрицательный ответ, следует признать, что простая прямая сборка in vitro может и не отвечать в точности процессу in vivo. Например, при сборке in vivo некоторых вирусов нужны вспомогательные белки, способствующие организации вирусной оболочки или упаковке ДНК в ней. Однако эти белки покидают вирион еще до окончания сборки.

Другим распространенным случаем является ферментативная модификация третичной или четвертичной структуры после укладки полипептидной цепи. Наглядным примером такого рода является инсулин. Выделенный из клетки, этот маленький белок состоит из двух полипептидных цепей, соединенных двумя дисульфидными мостиками (рис.1.11). Многочисленные попытки рекомбинировать эти две отдельные цепочки in vitro оказались

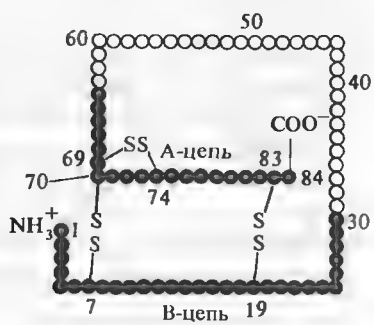


РИС. 1.11. Первичная структура проинсулина свиньи. Остатки от 31 до 63 вырезаются, в результате чего образуются А- и В-цепи функционально активного инсулина.

малоэффективными. Позже было обнаружено, что инсулин синтезируется in vivo как одна полипептидная цепь. После того как она сложится необходимым образом, большой центральный участок ее вырезается ферментативным путем, в результате чего в клетке образуется структура из двух цепочек. Когда сборка сопровождается такой химической модификацией, понимание ее механизма зависит от выяснения временной последовательности различных стадий укладки и модификации. Было бы очень важно попытаться определить, является ли та или иная модификация предварительным термодинамическим условием прохождения данного этапа укладки цепи или же она обусловлена просто кинетическими причинами.

ВОПРОСЫ, КАСАЮЩИЕСЯ ФУНКЦИИ

Знать структуру макромолекулы, не имея никакого представления о ее функции, — это то же самое, что любоваться зданием, не зная ничего о его назначении или его обитателях. Это может доставлять эстетическое удовольствие, но не дает пищи уму. Некоторые вопросы, касающиеся функции, относятся целиком к сфере биохимии или биологии, но можно назвать несколько проблем, требующих внимания со стороны биофизической химии.

1. Можно ли предсказать или объяснить функцию? Если задана только структура молекулы, было бы необыкновенным подвигом предсказать все разнообразие ее биологических функций, исходя только из общих знаний в области биохимии клетки и из структуры других молекул, не располагая дополнительной, полученной из опыта информацией об этой молекуле. Такой путь можно назвать молекулярной археологией, поскольку он сродни задаче воссоздания образа жизни и социальной структуры некоего общества по сохранившимся зданиям и другим предметам. Предположим, например, что расшифрована структура активного центра химотрипсина, показанная на рис. 1.10. Можно ли на этой основе предсказать, что химотрипсин является гидролитическим ферментом и что скорость осуществляемого им гидролиза сильно различается в случае сложных эфиров и амидов?

Вероятно, уместнее задать вопрос, какая дополнительная информация, кроме структурной, необходима для полного понимания функции. Вообще говоря, для того, чтобы предложить правдоподобный механизм работы фермента, помимо структурных, необходимы кинетические и термодинамические данные. Разница между знанием того, что процесс правдоподобен со структурной точки зрения, и утверждением, что он достаточно вероятен энергетически или кинетически, чтобы иметь биологическое значение, очень велика.

2. Все ли в структуре является необходимым? Этот вопрос относится к каждому уровню структурной организации и к каждой данной функции. Например, необходима ли для выполнения определенной функции интактная четвертичная структура или изолированные субъединицы могут выполнять работу столь же хорошо? Если это так, то какие именно субъединицы ответственны за эту функцию? Одним из наиболее интересных примеров, когда свойства отдельных субъединиц оказывается возможным сравнить со свойствами нативной четвертичной структуры, является белок триптофансинтаза *E. coli*. Он имеет субъединичную структуру $\alpha_2\beta_2$, и после диссоциации можно получить очищенные α- и β_2-субъединицы. Они катализируют следующие реакции:

$$\alpha\text{-Субъединица: Индол-3-глицерофосфат} \rightleftarrows \text{Индол} + \text{Глицеральдегид-3-фосфат}$$

$$\beta_2\text{-Субъединица: Индол} + \text{Серин} \rightleftarrows \text{Триптофан} + H_2O$$

В случае интактного $\alpha_2\beta_2$-фермента наблюдается следующая результирующая реакция:

$$\text{Индол-3-глицерофосфат} + \text{Серин} \rightleftarrows \text{Триптофан} + \text{Глицеральдегид-3-фосфат} + H_2O$$

На получение или использование в этой последней реакции свободного индола нет ника-

ких указаний. Это можно проверить, например, путем добавления в реакционную смесь радиоактивно меченного индола с тем, чтобы установить, включается ли он в триптофан.

Скорость синтеза триптофана интактным ферментом на порядок выше, чем этого следовало ожидать на основании свойств изолированных субъединиц. Может быть, это обусловлено просто тем, что индол не находится в равновесии с раствором? Чтобы проверить это, изучали мутантные формы фермента. Исходя из известной генетики этой системы, можно было получить белки, содержащие либо неактивные α-цепи, $\alpha_2^i\beta_2$, либо неактивные ß-цепи, $\alpha_2\beta_2^i$. Они осуществляли только частичные реакции, катализируемые изолированными активными субъедиинцами. Однако скорость катализа в случае $\alpha_2^i\beta_2$-белка была значительно выше, чем в случае β_2-субъединицы, а в случае $\alpha_2\beta_2^i$-белка она была выше, чем для изолированных α-субъединиц. Следовательно, каждая субъединица, входящая в иативную четвертичную структуру, как-то изменяет свойства другой, так что скорость катализа возрастает.

Какие особениости третичной структуры необходимы для выполнения данной функции и какие не важны для нее (если такие есть)? Вообще говоря, для функционирования большинство белков и нуклеиновых кислот должны иметь интактную третичную структуру. Сравнивая структуру нативных и денатурированных форм, иногда можно выделить те черты третичной структуры, которые необходимы для биологической активности. В случае полифункциональной молекулы часто удается выявить участки, играющие разную функциональную роль, применяя разделение ее на части путем последовательного разрыва связей. Биологи считают, что биологическую роль играет вся первичная структура. Это предположение можно считать достаточно обоснованным, в противном случае трудно было бы понять, почему в ходе эволюционного отбора несущественные части не были элиминированы и не прекратился их синтез. Биологическая роль некоторых участков молекулы может проявляться лишь косвенным путем. Например, некоторые из них могут оказаться необходимыми для того, чтобы облегчить соответствующую укладку остальных частей. In vitro, однако, удаление определенных участков макромолекулы иногда не приводит к нарушению ее функций. Это, вероятно, означает, что мы до сих пор не понимаем всех тонкостей поведения этой макромолекулы или тех факторов, которые in vivo регулируют ее синтез, активность и распад.

3. *С какими другими молекулами взаимодействует данная молекула?* Обычно первые указания на взаимодействие данной макромолекулы с какими-то другими из огромного множества малых и больших биологических молекул получают при исследовании ее функций. Затем пытаются установить, сколько молекул данного типа может взаимодействовать с рассматриваемой макромолекулой одновременно, как прочно они связываются и все ли связываются одинаково. На эти вопросы можно ответить с помощью прямых термодинамических измерений связывания (например, с помощью равновесного диализа) или применяя менее прямые спектроскопические методы. Если связывающиеся молекулы обладают подходящими спектроскопическими свойствами (например, если они окрашены), за ними наблюдают непосредственно. В противном случае иногда используют спектроскопические свойства молекулы-матрицы, с которой они связываются. В более строгих исследованиях изучают термодинамику и кинетику процесса связывания.

В случае связывания малых молекул часто можно производить систематическую модификацию их структуры, и это позволяет определить, какие их особенности узнаются белком или нуклеиновой кислотой и с какой целью. В качестве простого примера можно привести белок трансферрин, очень прочно связывающий Fe^{2+} в двух эквивалентных участках. Его сродство ко многим другим известным двухвалентным ионам, например к Ca^{2+} и Mg^{2+}, значительно ниже. Отсюда следует, что трансферрин — это белок, переносящий железо.

Если макромолекула специфически взаимодействует с другими молекулами разного вида, то желательно знать, как присутствие одних влияет на связывание других. Конкурент-

ное связывание означает, что в нем участвуют одни и те же (или перекрывающиеся) или же отдельные, но взаимно исключающие друг друга места; последнее может быть, например, обусловлено тем, что связывание одной молекулы вызывает такое структурное изменение в макромолекуле, которое ликвидирует место связывания другой молекулы. С другой стороны, связывание может быть кооперативным, при котором присутствие одной малой молекулы стимулирует связывание других либо того же сорта, либо совершенно другого. Эти эффекты могут быть весьма сложными, если в них участвует одновременно несколько разных типов взаимодействий. Ниже мы покажем на ряде примеров, как можно анализировать подобные случаи.

4. Какова структура функционального центра? Большинство функций макромолекул может быть связано с определенными их участками. Поэтому все вопросы о структуре, затронутые выше, приобретают решающее значение применительно к этим участкам. Кроме того, желательно знать, какие остатки в макромолекуле действительно контактируют со связанной молекулой. Расположены ли некоторые из них так, чтобы играть роль в катализе? Особенно важен вопрос о том, сопровождается ли связывание или функционирование структурными изменениями в «хозяйской» макромолекуле или в связанной с ней молекуле. Например, многие белки представляют собой смесь молекул, находящихся в нескольких дискретных конформациях, и лишь одна из них может оказаться стабильной при насыщении белка связывающимися с ним молекулами. Многие малые молекулы, связывающиеся с двойными спиралями нуклеиновых кислот, вызывают некоторое их раскручивание, удлиняя всю структуру. Все это означает, что далеко не всегда структура пассивной, не работающей макромолекулы идентична структуре функционирующей макромолекулы.

Если при функционировании происходят структурные изменения, прежде всего необходимо выяснить, на каком структурном уровне они осуществляются и какие участки в них вовлечены. Можно подумать, что в случае окси- и дезоксигемоглобина (рис. 1.5) главное структурное изменение, сопровождающее оксигенацию, состоит в перестройке четвертичной структуры, но это не обязательно так. Кислород связывается с группами гема, расположенными вблизи участков контактирования четырех субъединиц белка. Но для того, чтобы произошло изменение четвертичной структуры, свойства остатков на поверхности должны быть как-то изменены. Действительно, когда кислород связывается с гемом гемоглобина, атом железа в геме сдвигается, что вызывает ряд небольших изменений третичной структуры, изменяющих поверхность субъединиц. Эти изменения не менее важны для понимания механизма кооперативного связывания кислорода гемоглобином, чем значительно более заметные изменения четвертичной структуры. Иногда в макромолекуле не происходит структурных перестроек при связывании с ней другой молекулы, но зато последняя претерпевает такие изменения. Прекрасным примером такого рода является связывание гексасахарида с молекулой фермента лизоцима. Как показано на рис. 1.12, один из участков связывания сахара в молекуле лизоцима не способен присоединять сахар в нормальной конформации кресла. Для того чтобы произошло такое связывание, сахарное кольцо должно деформироваться и перейти в форму полукресла, что энергетически не выгодно. Тем не менее этот переход осуществляется, так как затрата энергии с лихвой компенсируется энергией связывания остальных молекул сахара. Грубо говоря, лизоцим способен использовать энергию связывания, сконцентрировав ее в одной точке углеводного комплекса. Это помогает разрыву связи С—О в одном из сахаров, что является частью механизма каталитического действия фермента.

В случае лизоцима структурные изменения в области активного центра достаточно велики, и благодаря этому их можно четко видеть. Однако способствовать протеканию химических реакций в активном центре могут и значнтельно меньшие изменения. Таковы, например, изменения детальных особенностей локального окружения. Главной целью биофизической химии является количественное объяснение эффективности ферментов как

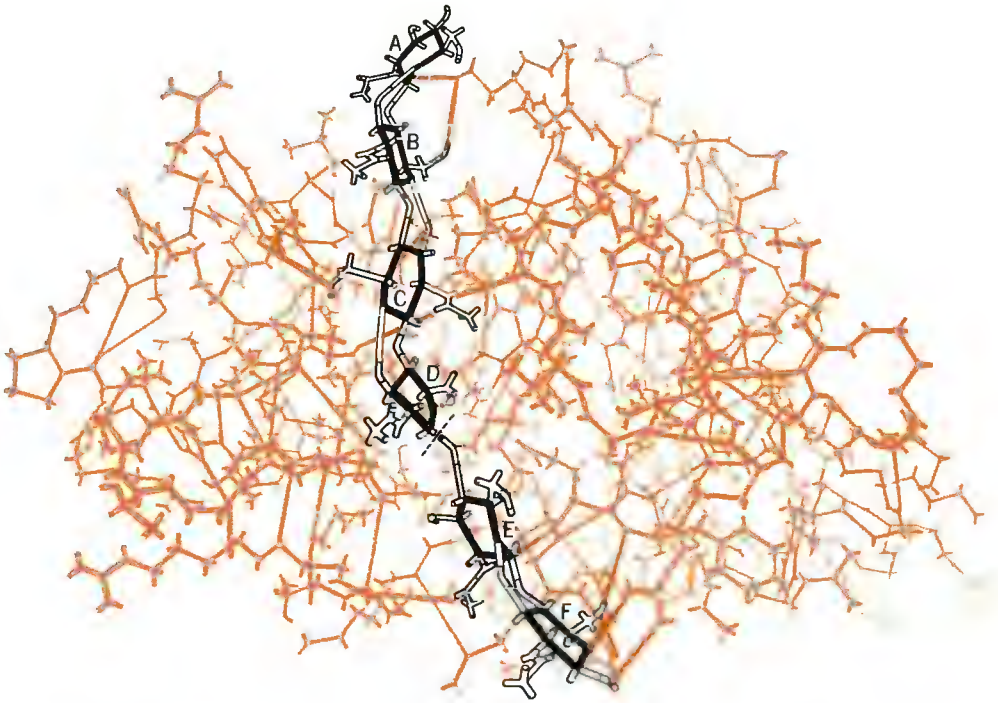


РИС. 1.12. Третичная структура лизоцима белка куриного яйца в комплексе со связанным субстратом $(NAG—NAM)_3$. Химическую структуру субстрата см. на рис. 1.2. Ярко окрашенные палочки — водородные связи между ферментом и субстратом. Отметим, что сахар, обозначенный буквой D, деформирован так, что четыре атома лежат в одной плоскости. Место разрыва связи под действием лизоцима отмечено пунктиром. (Рисунок Ирвинга Гейса.)

катализаторов. Одна из трудностей здесь состоит в том, что эффект возникает в результате совместного действия большого числа факторов, каждый из которых сам по себе весьма незначителен и поэтому с трудом поддается измерению.

5. Как осуществляется регуляция функции? Для большинства процессов, протекающих в живых организмах, характерен феномен гомеостаза, т.е. в организме наблюдается тенденция сопротивляться любым изменениям или компенсировать их. Способность биологических молекул к саморепликации и катализу обширных классов реакций таит в себе опасность возникновения химической нестабильности. В конце концов, взрыв является просто автокаталитической реакцией. Ясно, что живые организмы научились контролировать свои химические процессы так, чтобы обеспечить правильный баланс синтезируемых веществ, эффективное использование имеющейся энергии и сглаживание флуктуаций широкого круга параметров окружающей среды. Это означает, что функции многих биологнческих молекул должны регулироваться в ответ на нужды организма. Такая регуляция может происходить на различных уровнях.

Рассмотрим возможные механизмы регуляции активности фермента в клетке. Одним из регулируемых параметров является общее количество фермента. Эта величина определяется скоростями синтеза и распада. Другой параметр — доля активного фермента. Она может изменяться, например, при изменении равновесия на уровне четвертичной структуры, образующейся из активных и неактивных форм. Такое равновесие может регулиро-

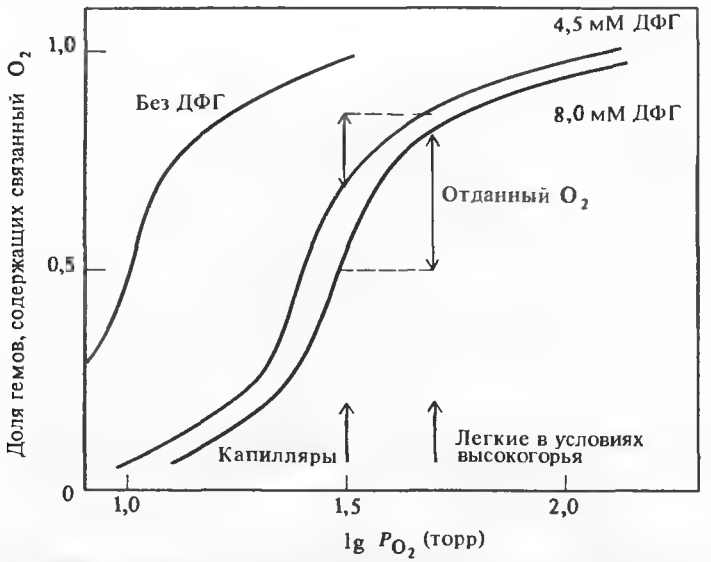


РИС. 1.13. Связывание кислорода гемоглобином как функция парциального давления кислорода. Стрелками отмечено примерное парциальное давление в капиллярах (30 торр) и в легких в условиях высокогорья (50 торр). Приведены кривые связывания для трех различных концентраций дифосфоглицерата. Способность гемоглобина отдавать кислород определяется разностью количества связанного кислорода в легких и в капиллярах. Из рисунка видно, как эта разность возрастает при высоком уровне дифосфоглицерата.

ваться эффекторами — молекулами, которые не обязательно участвуют в реакции, катализируемой ферментом, но просто сдвигают ее в сторону образования активных или неактивных форм.

Активность фермента может регулироваться также путем изменения числа активных центров за счет заполнения некоторых из них молекулами ингибитора. Одним из способов регуляции является изменение прочности связывания молекул, подвергающихся каталитической реакции, путем изменения скорости их связывания или освобождения или благодаря изменению собственной скорости катализа, осуществляемого ферментом. Эффективная активность может регулироваться и на более высоком уровне, путем изменения количества молекул субстрата, на который действует фермент, или путем изменения их локализации в клетке. От этого зависит, смогут ли молекулы субстрата оказаться в сфере действия фермента.

Точно такого же рода рассмотрение приложимо к тем белкам и нуклеиновым кислотам, которые не участвуют в катализе. Их функции также регулируются на различных уровнях. Прекрасным примером является гемоглобин. Мы отметили выше, что дифосфоглицерат изменяет свойства гемоглобина: связываясь с дезокси-формой гемоглобина (рис. 1.5), он стабилизирует ее относительно окси-формы, в результате чего при заданной концентрации кислорода его связывание уменьшается. Влияние изменения содержания дифосфоглицерата на связывание гемоглобином кислорода показано на рис. 1.13. Гемоглобин доставляет кислород из легких в капиллярную сеть всех тканей. Количество кислорода, которое гемоглобин связывает в легких, определяется парциальным давлением кислорода в воздухе. На уровне моря оно равно примерно 100 торр. В капиллярах гемоглобин освобождает кислород, и устанавливается новое равновесие, определяющееся уровнем кислорода в тканях, приблизительно соответствующим давлению 30 торр. Разность двух уровней связывания — это количество кислорода, которое каждая молекула гемоглобина может доставить в ткани.

Дифосфоглицерат сдвигает кривую связывания кислорода гемоглобином эритроцита. Форма этой кривой такова, что увеличение концентрации дифосфоглицерата приводит к более эффективному освобождению кислорода при более низком его давлении. Поэтому одной из немедленных реакций организма на уменьшение атмосферного давления (например, с высотой) является повышение уровня дифосфоглицерата. Это — срочная мера. Долговременное решение заключается в синтезе большего количества гемоглобина и в увеличении числа эритроцитов. Так и происходит, но это значительно более медленный процесс, требующий нескольких дней, в то время как дифосфоглицератная адаптация происходит за часы.

В случае молекул, выполняющих более чем одну функцию, процесс регуляции может стать весьма сложным. Могут ли разные функции осуществляться одновременно? Если да, то возможна ли их независимая регуляция? Если нет, то чем определяется основная функция? На практике обычно изучают регуляцию активиости отдельных видов макромолекул. Однако если для выполнения какой-то функции необходимо согласованное действие целого набора молекул, ясно, что должен существовать способ не только управления функционированием отдельных видов молекул, но и регулирования всего ансамбля в целом. Эта задача, по-видимому, решается с помощью более общих регуляторных процессов, использующих большие и малые молекулы, предназначенные и синтезируемые исключительно для целей регуляции. Сюда относятся стероидные и пептидные гормоны, циклические нуклеотиды, а также белки, которые их синтезируют или взаимодействуют с ними. Такие белки являются одним из наиболее интересных объектов биофизических исследований, поскольку они играют центральную роль в биологических процессах.

В функции других больших молекул входит поддержание структуры клетки и контроль переноса веществ в клетку и из нее. Исследование регуляции этих функций представляется особенно привлекательным, так как может оказаться необходимой координация на молекулярном уровне событий, происходящих на большом расстоянии друг от друга.

1.3. Некоторые стратегические подходы, используемые в биофизической химии

Получение ответа на многие поставленные нами вопросы связано с большими трудностями. Эти трудности были бы очень велики даже в том случае, если бы речь шла только о малых молекулах. Сложность же крупных молекулярных образований накладывает жесткие ограничения на возможные виды экспериментов и делает любой эксперимент весьма трудоемким. Наиболее серьезным препятствием является огромное число параметров, определяющих структуру макромолекул, и множество параметров окружающей среды, потенциально связанных с их функцией. Большое число переменных ведет к двум неприятным последствиям. Число возможных экспериментов растет с числом переменных в геометрической прогрессии, поэтому вероятность случайного выбора действительно важного эксперимента безнадежно мала. Вторая проблема состоит в том, что даже если сделано очень много опытов, трудно осмыслить их результат с единой точки зрения.

Большинство исследователей считают более удобным рассматривать реакцию на одну переменную; результаты в этом случае могут быть представлены графически в двумерном пространстве. Можно также иметь дело с двумя переменными одновременно, пользуясь изображением поверхности в трехмерном пространстве. Но почти всегда приходится ограничиваться одновременным рассмотрением не более двух переменных. В то же время структура белка может описываться сотней важных переменных, и мало кто способен удержать в памяти эти многочисленные данные и использовать их в нужный момент для проведения необходимых корреляций. Поэтому для подробных биофизических исследований необходимо упрощать систему. Стратегия здесь отвечает известному принципу — разделяй и властвуй. Необходимо выделить часть структуры, одиу из пере-

менных, одно явление, а остальное контролировать или пытаться обеспечить его постоянство, так чтобы одиовременно анализировать и пытаться понять лишь один аспект проблемы. Имеется несколько главных путей реализации этой стратегии.

ИСПОЛЬЗОВАНИЕ СИСТЕМЫ МЕНЬШИХ РАЗМЕРОВ В КАЧЕСТВЕ МОДЕЛИ

Структурное разрешение любого метода ухудшается с ростом размеров молекулы. В принципе нет причин, по которым нельзя было бы определить структуру большой системы с разрешением в 1 Å или еще лучшим, но существуют огромные трудности практического характера (см. гл. 13). Гораздо легче исследовать небольшую часть системы с очень высоким разрешением, а затем использовать эти результаты для того, чтобы судить о ее структуре в целой системе. Например, трудно получить точные координаты атомов α-спиральных участков виутри глобулярного белка или атомов двойных спиралей РНК внутри ее третичной структуры. Поэтому имеет смысл исследовать малые модельные спиральные структуры в качестве структурных моделей. Далее необходимо найти способ убедиться в том, что установленная таким образом структура существенно не изменяется, входя как часть в молекулу большего размера. В качестве другого примера можно привести спектроскопические методы, являющиеся очень мощным орудием структурного исследования в случае достаточно малых молекул, у которых спектральный вклад каждой группы может быть определен и измерен по отдельности. Поскольку спектры молекул большего размера получаются путем усреднения по многим звеньям, детальная их интерпретация затрудиительна. Если, однако, свойства отдельных звеньев в заданном окружении и в заданной конформации известны из модельных исследований, то иногда удается установить соответствие спектра полимерной системы определенным гипотезам относительно ее структуры.

НАБЛЮДЕНИЕ ЛИШЬ ЗА ЧАСТЬЮ СИСТЕМЫ

Работа с частью системы, физически отделенной от системы в целом, всегда таит в себе опасность получения неверных результатов, поскольку эта часть необязательно будет вести себя так же, как в интактной системе. Одна из возможностей обойти эту трудность заключается в том, чтобы найти способ наблюдать одновременно лишь за этой одной частью, но в составе интактной системы. Это основная цель исследований, в которых используются зонды или метки. Иногда макромолекула сама содержит такой зонд, например ион переходного металла в функциональном центре или вблизи него или окрашенную простетическую группу, скажем гем. В других случаях исследователь вводит зонд в заданное место структуры путем его ковалентного присоединения или нековалентного связывания. От вида используемого зонда зависит, какие локальные особенности удастся выявить. К этим особенностям относятся полярность, гибкость, присутствие определенных звеньев макромолекулы или других связанных молекул, истинная конформация самого зонда, его электронная структура и природа его связи с макромолекулой.

Все исследования, опирающиеся на введение зонда, связаны с определенным риском. Зонд позволяет «заглянуть внутрь» системы, но ценой возможного нарушения ее структуры. Это нарушение скорее всего произойдет в области, непосредственно примыкающей к зонду. К сожалению, это и есть та область, о которой мы хотим получить сведения с помощью зонда.

СРАВНЕНИЕ МЕЖДУ СОБОЙ ПОЧТИ ИДЕНТИЧНЫХ СИСТЕМ

Прямой подход к выявлению роли отдельных структурных переменных состоит в том, чтобы производить в структуре малые возмущения и наблюдать их результаты. Как

это делается? Один подход заключается в сравнении свойств молекул, очень мало отличающихся друг от друга. Например, некоторые белки (цитрохром *c*, гемоглобин и др.) и РНК (например, транспортная РНК) были выделены из разных источников. Поскольку структура и функциональные свойства у них почти одинаковы, это позволяет судить о том, какие части структуры должны сохраняться неизменными. По-видимому, именно эти области необходимы для функционирования. Более информативным подходом может являться сравнение структур, сходных между собой, но различающихся функционально. В этом случае иногда удается определить ключевые структурные особенности. Например, для тРНК, участвующих в синтезе пептидогликана клеточной стенки, характерны определенные четкие изменения структуры по сравнению с подобными молекулами, участвующими в белковом синтезе на рибосоме (см. гл. 3). Многие белки имеют гомологичную структуру, но разные функции. Одним из примеров являются сериновые эстеразы — химотрипсин, трипсин, эластаза и т.д. Бывают даже случаи, когда молекулы, например пероксиддисмутаза и часть иммуноглобулина G, абсолютно несхожие функционально (насколько мы можем судить), имеют тем не менее аналогичную структуру.

Особенно эффективной стратегией выделения структурных и функциональных переменных является использование мутантных белков или нуклеиновых кислот. Преимущество этого подхода заключается в том, что в таких молекулах произошли замены лишь одной или нескольких аминокислот (или соответственно нуклеотидов), но это привело к таким функциональным изменениям, что позволило идентифицировать объект исследования как мутантный. В случае бактерий здесь разработаны чрезвычайно искусные приемы, с помощью которых избирательно создают изменения в нужном белке или нуклеиновой кислоте, а затем побуждают организм усилить синтез мутантных продуктов в такой степени, чтобы можно было выделить их в препаративных количествах. Именно таким путем были получены сотни мутантных белков-репрессоров лактозного оперона, содержащих отдельные аминокислотные замены. Теперь стало возможным с помощью тонких генетических манипуляций изменять по желанию почти любую аминокислоту. Это непростая задача, но все же этот путь проще, чем органический синтез, поскольку после того, как создан нужный штамм, он будет всегда производить мутантный белок.

В случае клеток эукариот техника генетических манипуляций разработана значительно слабее. Здесь больше полагаются на спонтанные мутации, приводящие к какому-либо фенотипическому дефекту. Последний может быть обнаружен непосредственно в полевых исследованиях, в лаборатории или в клинике, иногда в результате систематического поиска, но часто как побочный продукт другой работы. Например, многие мутации, затрагивающие гемоглобин, приводят к анемиям, а последние легко распознаются клинически. В результате известно очень большое число мутантных форм гемоглобина человека, установлены различия в аминокислотной последовательности, а в некоторых случаях поняты причины функциональных изменений. Например, гемоглобин Сидней содержит единственную замену — валин в 67-м положении β-цепи заменен в нем на аланин. Фенотипическим проявлением такой замены в клинике является гемолитическая анемия. Эритроциты при этом становятся хрупкими и легко разрушаются. Исследования выделенного из них гемоглобина показывают, что гем β-цепей слабо связан с белком и легко отделяется. Это приводит к двум последствиям. Дефицитный по гему белок становится значительно менее эффективным переносчиком кислорода — как потому, что содержит на два связывающих кислород гема меньше, чем нормальный, так и вследствие изменения формы кривой связывания кислорода, приводящего к менее эффективному освобождению кислорода в капиллярах. Кроме того, отделяющийся гем вреден: он, по-видимому, взаимодействует с клеточной мембраной эритроцита, что каким-то образом увеличивает ломкость последнего. Непрочности связывания гема с гемоглобином Сидней можно дать простое объяснение на молекулярном уровне. В нормальном гемоглобине валин способствует образова-

нию «кармана», в котором и происходит связывание гема. Сродство к гему здесь частично обусловлено стремлением гема избежать контакта с водой, расположившись в неполярном окружении кармана. В мутантном гемоглобине, когда валин заменен на аланин, карман оказывается слишком мал для того, чтобы в нем мог разместиться гем, и свободное пространство, по-видимому, заполнено водой. Это, естественно, ослабляет неполярные взаимодействия.

Когда генетические вариации отсутствуют или не удовлетворяют требованиям опыта, иногда применяется избирательная химическая модификация. Некоторые аминокислотные остатки могут быть избирательно модифицированы в присутствии остальных. Даже если в белке имеется много экземпляров остатков каждого типа, часто лишь ограниченное их число обладает высокой реакционной способностью. Если в результате модификации структура или функция белка изменяются, можно попытаться объяснить, каким образом эта модификация привела к такому результату. В случае нуклеиновых кислот осуществление специфической химической модификации затруднительно, поскольку здесь гораздо меньше видов звеньев, а химические свойства всех звеньев, к сожалению, очень близки. И все же существуют приемы, способные упростить задачу. Предположим, что в результате модификации получается гетерогенная смесь молекул, прореагировавших в нескольких различных местах. Вероятно, некоторые из этих молекул будут инактивированы, а другие сохранят нормальную функцию. Тогда можно будет с помощью физических методов, основанных на способности молекул к функционированию (например, с помощью аффинной хроматографии), разделить полученную смесь. На практике получение химически модифицированных макромолекул является нередко трудоемкой процедурой, проводимой методом проб и ошибок. Тем не менее использование таких макромолекул необходимо для ряда других методов, таких, как метод зондов или меток, а также для изоморфного замещения тяжелыми атомами при изучении больших молекул методом рентгеновской кристаллографии.

ВЫДЕЛЕНИЕ ДИСКРЕТНЫХ СОСТОЯНИЙ СИСТЕМЫ

Многие биологические полимеры могут существовать во множестве различных конформаций, которые стабилизируются по-разному под действием таких параметров среды, как температура, pH и концентрация соли. При детальных структурных исследованиях обычно необходимо работать одновременно с одной чистой формой. Поэтому бывает полезно рассмотреть разнообразные условия, чтобы определить области устойчивости различных конформаций и районы конформационных переходов. Поскольку часто структурные превращения молекул происходят в очень узком интервале условий, каждое чистое состояние можно уподобить фазе системы, а конформационные превращения называть фазовыми переходами. Все известные сведения относительно влияния внешних условий могут быть представлены в виде фазовой диаграммы, подобно тому как изображаются состояния воды или сплавов в зависимости от термодинамических переменных. Некоторые примеры построения этих диаграмм приведены в гл. 22 и 24. Важно, что после того, как такая фазовая диаграмма построена, из нее сразу видно, какие пределы внешних условий можно использовать для изучения структуры чистых состояний и в каких областях следует проводить исследование конформационных переходов. Это может существенно упростить дальнейшие исследования.

Если конформация изменяется постепенно, в широком интервале внешних условий, трудно построить фазовую диаграмму. Более того, становится неясно, описывать ли систему как ряд дискретных состояний, постепенно переходящих друг в друга, или предположить, что имеется ансамбль из многих состояний.

Термодинамические и кинетические исследования, которые могут помочь разрешить этот вопрос, описаны в гл. 20 и 21. Обычная тактика сводится к попытке рассмотрения

системы как смеси минимального количества дискретных состояний, необходимого для описания имеющихся экспериментальных данных. В простых случаях это число часто определяется прямо из рассмотрения системы. Для сложных случаев имеется математический аппарат, позволяющий определить число переменных, необходимое для описания населенности дискретных состояний в пределах точности измерений (Дополнение 1.1).

Дополнение 1.1

ОБЩАЯ СХЕМА МНОГОКОМПОНЕНТНОГО АНАЛИЗА

Предположим, что измерены l различных экспериментальных параметров, каждый в r экспериментальных условиях. Обозначим через P_{ji} значение i-го параметра в j-х условиях опыта. Совокупность P_{ji} можно представить в виде матрицы $\underset{\sim}{\mathbf{P}}$ размером $l \times r$. Можно далее построить симметричную ($r \times r$) квадратную матрицу $\underset{\sim}{\mathbf{Q}} = \underset{\sim}{\mathbf{P}}\underset{\sim}{\mathbf{P}}^{+}$ умножением матрицы $\underset{\sim}{\mathbf{P}}$ на траиспонированиую ($P^{+}_{ij} = P_{ji}$). В матрице $\underset{\sim}{\mathbf{Q}}$ содержится вся известная нам информация о системе. Можно обычными методами найти собственные числа λ_k и собственные векторы Ψ_k матрицы $\underset{\sim}{\mathbf{Q}}$. Если $\underset{\sim}{\mathbf{E}}$ — единичная матрица, то

$$|\underset{\sim}{\mathbf{Q}} - \lambda_k \underset{\sim}{\mathbf{E}}| = 0 \text{ и } \underset{\sim}{\mathbf{Q}}\Psi_k = \lambda_k \Psi_k \text{ где } k = 1, \ldots, r$$

Оказывается, что векторы Ψ_k образуют базисный набор для описания системы, аналогично тому как волновые функции гамильтониана молекулы могут представлять состояния молекулы. Собственные числа λ_k оказываются мерой вклада данного базисного состояиия. На практике можно, просмотрев все λ_k, ограничиться для описания системы лишь наибольшими из них. Для решения вопроса о том, сколько соответствующих собственных векторов Ψ_k следует включить в описание свойств системы, можно использовать статистические тесты. Такая процедура является очень формальной, но достаточно общей. На практике, однако, экспериментальные данные редко обладают достаточной точностью для оправдания ее использования, и следует применять приближенное рассмотрение. Пример использования этой схемы многокомпонентного анализа можно найти в работе R. H. Fairclough et al., J. Mol. Biol., 132, 557 (1979).

Краткие выводы

Имеются четыре уровня структурной организации биологических молекул. Первичная структура — это ковалентная химическая структура или последовательность звеньев в цепи полимера плюс все поперечные сшивки внутри цепей или между ними. Вторичная структура — это локально упорядоченное спиральное расположение звеньев. Третичная структура — это полная трехмерная структура единицы, которая не может быть разделена без нарушения ковалентных связей. Четвертичная структура — это расположение отдельных единиц третичной структура внутри комплекса.

Среди иаиболее важных вопросов, рассматриваемых в биофизической химии, отметим следующие. Каким образом можно установить структуру? Можно ли ее предсказать или разумно объясиить? Следует ли считать структуры статическими или динамическими? Как упорядоченность изменяет физические и химические свойства молекулы? Как связаны структура и биологическая функция? Каков механизм регуляции биологической функции?

Основные стратегические подходы к исследоваиию биологических полимеров основаны на уменьшении числа используемых переменных. Обычно бывает необходимо упро-

стить систему либо путем использования в качестве модели системы меньших размеров, либо применением методов, позволяющих наблюдать за частью системы, либо путем сравнения двух почти идентичных систем, когда могут быть поняты эффекты, обусловленные малыми различиями систем. Наконец, та же цель может быть достигнута путем выделения дискретных состояний системы подбором внешних условий. Большинство этих подходов может быть кратко определено как «разделяй и властвуй».

ЗАДАЧИ

1.1. Имеется непрерывная спираль радиуса r с шагом P. Покажите, что путь, проходимый вдоль одного витка спирали (контурная длина), равен $\sqrt{P^2 + 4\pi^2 r^2}$. *Указание:* спроецируйте спираль на цилиндрическую поверхность, на которой она лежит, а затем разверните цилиндр.

1.2. Предположим, что некий белок способен регулировать свой собственный биосинтез. В растворе белок может существовать в двух формах — мономерной и октамерной, между которыми существует равновесие. Мономер нестабилен и может денатурировать. Он стимулирует синтез, повышая его скорость; октамер ингибирует синтез. Объясните, почему это оказывается выгодным для организма.

ЛИТЕРАТУРА

Общая

Bloomfield V. A., Harrington R. F., eds. 1975. Biophysical Chemistry, San Francisco, W. H. Freeman and Company. [Перепечатка соответствующих статей из Scientific American.]
Stryer L., 1975. Biochemistry, San Francisco, W. H. Freeman and Company. [Содержит обзорные данные о биохимической роли различных макромолекул.]

Специальная

Dayhoff M. A., ed., 1972-1976. Atlas of Protein Sequence and Structure, Silver Spring, Md., National Biomedical Research Foundation.
Eisen H. N., 1974. Immunology, New York, Harper and Row. [В гл. 15 и 16 рассмотреиы структура антител и взаимодействия антиген — антитело.]
Kirschner K., Bisswanger H., 1976. Multifunctional proteins, Ann. Rev. Biochem., 45, 143.
Kraut J., 1977. Serine proteases: Structure and mechanism of catalysis, Ann. Rev. Biochem., 46, 331.
Magar M. E., 1972. Data Analysis in Molecular Biology, New York, Academic Press. [Рассматриваются различиые вопросы многокомпонентного анализа.]
Nomura M., Tissieres A., Lengyel P., 1974. The Ribosome, Long Island, N. Y., Cold Spring Harbor Press. [Наиболее полная работа по этому вопросу.]
Perham R. N., 1975. Self-assembly of biological molecules, Phil. Trans. Roy. Soc. London Ser. B, 272, 123. [Очень интересиая, иебольшая по объему статья.]
Pogson C. J., 1974. The protein chemistry of regulatory phenomena. In: Biochemistry Series One, 1: Chemistry of Macromolecules, ed. H. Gutfreund, London, Butterworths, p. 287.

2.1. Свойства аминокислот

Элементарными единицами, из которых состоят белки, являются 20 α-аминокислот. Их структура, название и общепринятые трехбуквенные обозначения приведены на рис. 2.1. Во всех 20 аминокислотах, кроме глицина, α-углеродный атом асимметричен, причем в белках и пептидах, синтезируемых в рибосомах, встречаются только L-стереоизомеры (см. рис. 2.1), поэтому приставку L мы будем в дальнейшем опускать. Правда, в некоторых природных пептидах обнаружены D-аминокислоты, однако эти пептиды синтезируются не в рибосомах.

Две аминокислоты, изолейцин и треонин, имеют второй асимметричный центр при β-углеродном атоме. Их конфигурация тоже постоянна (см. рис. 2.1). Аминокислоты с другой конфигурацией при β-углеродном атоме называются аллоизолейцином и аллотреонином.

Первичной ковалентной структурой белка является полипептидная цепь:

$$NH_3^+—CHR_1—CO—NH—CHR_2—CO—\ldots—$$
$$—NH—CHR_{n-1}—CO—NH—CHR_n—COO^-.$$

Обычно, если это специально не оговорено, формулы пептидных цепей записывают таким образом, чтобы NH-группы во всех пептидных связях располагались справа. Например, запись Ile—Glu обозначает дипептид, в котором изолейцин — N-концевой остаток, а глутамин — C-концевой остаток.

ИОНИЗАЦИЯ АМИНО- И КАРБОКСИЛЬНЫХ ГРУПП В АМИНОКИСЛОТАХ И ПЕПТИДАХ

На рис. 2.1 приведены структурные формулы преобладающих при pH 7 ионизованных форм аминокислот. При таком значении pH α-аминогруппа протонирована, а карбоксильная группа депротонирована. Таким образом, при нейтральных pH аминокислоты не несут заряда, а существуют в виде цвиттерионов $NH_3^+—CHR—COO^-$, обладающих значительным дипольным моментом. Четыре аминокислоты при pH 7 имеют дополнительные заряды на боковой группе: аспарагиновая и глутаминовая кислоты — отрицательный, лизин и аргинин — положительный.

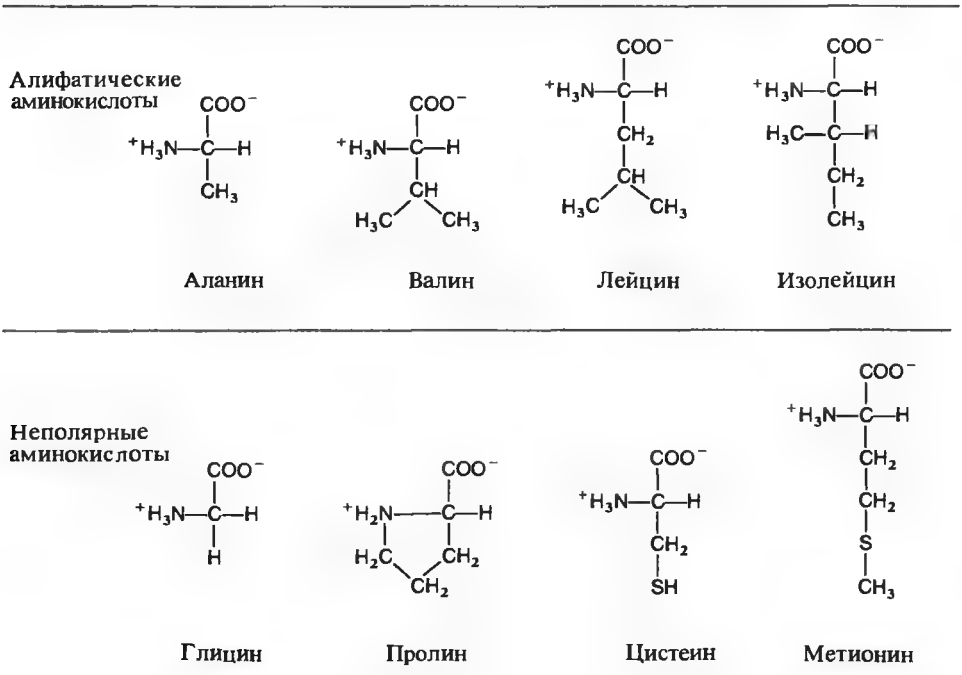
Так как свойства белков сильно зависят от pH среды, необходимо знать, при каких значениях этого параметра титруются различные функциональные группы аминокислот, а также то, как эти значения изменяются при включении аминокислот в белок. Для отдельных аминокислот и небольших пептидов такую информацию можно получить прямым потенциометрическим титрованием. Присоединение или потеря протона аминокислотой выявляются по изменению pH раствора. Для интактных белков полное потенциометрическое титрование при наличии достаточно большого количества вещества представляет собой довольно простую процедуру. Однако интерпретация результатов затруднена тем, что часто неясно, какой остаток из множества сходных или идентичных определяет ту или иную область кривой титрования. В таких случаях необходимо использовать другие, более специфичные методы исследования. Так, при титровании аминокислотных остатков с такими боковыми группами, как у гистидина, удобно применять ядерный магнит-

ный резонанс (ЯМР), в случае тирозина и подобных ему остатков весьма полезен спектральный анализ в оптическом диапазоне (см. гл. 7 и 9).

При образовании пептидных связей в белках и полипептидах отдельные аминокислотные остатки теряют свой цвиттерионный характер. Их кислотно-основные свойства определяются лишь двумя концевыми остатками и боковыми группами. При этом изменение кривой титрования на участках, соответствующих отдельным остаткам, объясняется в основном двумя эффектами: во-первых, электростатическими взаимодействиями, вызванными изменениями локальной химической структуры, и, во-вторых, термодинамическими взаимодействиями, обусловленными сопряжением ионного и конформационного равновесий. Рассмотрим разницу между N- и C-концами белков, с одной стороны, и амино- и карбоксильными группами свободных аминокислот — с другой. При pH 7 в свободных аминокислотах имеет место значительное электростатическое притяжение между положительно заряженной NH_3^+-группой и отрицательно заряженной COO^--группой. Данные по ионному равновесию олигоаланинов, приведенные в табл. 2.1, показывают, что это затрудняет удаление протона от NH_3^+, а также присоединение протона к COO^-. Таким образом, если для Ala или $(Ala)_2$ влияние заряженных групп очень велико, то уже для $(Ala)_4$ взаимодействие между концевыми заряженными группами не столь важно. В Дополнении 2.1 рассмотрены основные принципы ионизационного равновесия и проиллюстрировано использование данных титрования для оценки энергии электростатического взаимодействия.

Потенциометрическое титрование не позволяет прямо определить, какой именно остаток титруется. Такую информацию обычно можно получить либо путем спектроскопических исследований, либо путем использования химических аналогов изучаемых соеди-

РИС. 2.1. Структура 20 обычных аминокислот при pH 7. Все они изображены в плоскости α-углеродного атома. В действительности же аминогруппа и водород находятся над плоскостью рисунка, а COO-группа — под ней.

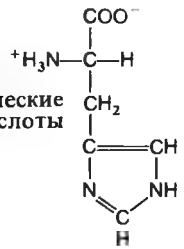


нений. Например, приведенные ниже аналоги дают возможность достаточно достоверно определить значение pK_a олигоаланинов:

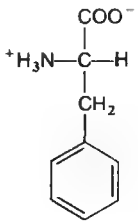
$$CH_3—CO—NH—CH(CH_3)—COOH \overset{1}{\rightleftarrows} CH_3—CO—NH—CH(CH_3)—COO^- \quad pK_1 = 3{,}72$$

$$NH_3^+—CH(CH_3)—CO—NH_2 \overset{2}{\rightleftarrows} NH_2—CH(CH_3)—CO—NH_2 \quad pK_2 = 8{,}02$$

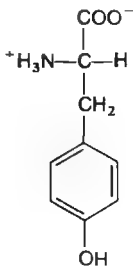
Ароматические аминокислоты



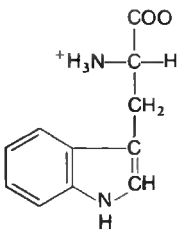
Гистидин



Фенилаланин

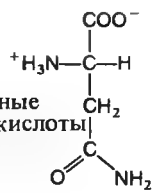


Тирозин

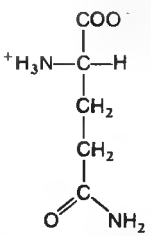


Триптофан

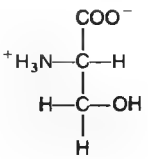
Полярные аминокислоты



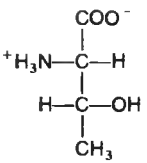
Аспарагин



Глутамин

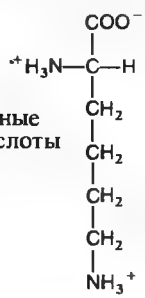


Серин

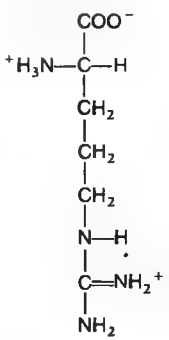


Треонин

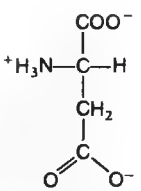
Заряженные аминокислоты



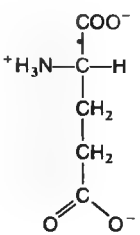
Лизин



Аргинин



Аспарагиновая кислота



Глутаминовая кислота

Таблица 2.1

ИОННОЕ РАВНОВЕСИЕ ДЛЯ НЕКОТОРЫХ ОЛИГОАЛАНИНОВ

Равновесие	pK_1	pK_2
$NH_3^+-CH(CH_3)-COOH \overset{1}{\rightleftarrows} NH_3^+-CH(CH_3)-COO^- \overset{2}{\rightleftarrows} NH_2-CH(CH_3)-COO^-$ (+Ala; +Ala; Ala)	2,34	9,69
$NH_3^+-CH(CH_3)-CO-NH-CH(CH_3)-COOH \overset{1}{\rightleftarrows}$ (+Ala—Ala) $\rightleftarrows NH_3^+-CH(CH_3)-CO-NH-CH(CH_3)-COO^- \overset{2}{\rightleftarrows}$ (+Ala—Ala) $\rightleftarrows NH_2-CH(CH_3)-CO-NH-CH(CH_3)-COO^-$ (Ala—Ala)	3,12	8,30
$^+Ala-Ala-Ala \overset{1}{\rightleftarrows} {}^+Ala-Ala-Ala^- \overset{2}{\rightleftarrows} Ala-Ala-Ala^-$	3,39	8,03
$^+Ala-Ala-Ala-Ala \overset{1}{\rightleftarrows} {}^+Ala-Ala-Ala-Ala^- \overset{2}{\rightleftarrows} Ala-Ala-Ala-Ala^-$	3,42	7,94

Дополнение 2.1

ПРИНЦИПЫ ИОНИЗАЦИОННОГО РАВНОВЕСИЯ

Реакция ионизации $H-A \rightleftarrows H^+ + A^-$ характеризуется константой диссоциации

$$K_a = (H^+)(A^-)/(H-A)$$

Удобно представить K_a в форме, аналогичной pH:

$$pK_a = -\lg K_a$$

В условиях, когда кислота ионизована наполовину, $(H-A) = (A^-)$, $K_a = (H^+)$, или $pK_a = pH$. Коистанта диссоциации связана со стандартной свободной энергией диссоциации следующим образом:

$$\Delta G^\circ = -RT \ln K_a$$

ΔG° есть разность свободной энергии образования продуктов реакции (H^+ и A^-) и реагента ($H-A$), когда они находятся в стандартном состоянии (допустим, 1 М водный раствор). Действительное изменение свободной энергии при ионизации, происходящей в других условиях, равно

$$\Delta G_{\text{иониз}} = \Delta G^\circ + RT \ln [(H^+)(A^-)/(H-A)]$$

Таким образом, если концентрации (H^+), (A^-) и $(H-A)$ являются равновесными [т.е. $(H^+)(A^-)/(H-A) = K_a$], то $\Delta G_{\text{иониз}} = 0$.

Допустим, что реакция ионизации сопряжена с каким-либо другим процессом или взаимодействием, например присоединение протона к A^- изменяет взаимодействие A^- с некоторыми другими группами молекулы. Обозначим измеиение свободной энергии, обусловленное этим взаимодействнем, через $\Delta G_{\text{сопр}}$. Тогда полная свободная энергия в реакции ионизации будет равна

$$\Delta G_{\text{поли}} = \Delta G_{\text{иониз}} + \Delta G_{\text{сопр}} = \Delta G^\circ + \Delta G_{\text{сопр}} + RT \ln [(H^+)(A^-)/(H-A)]$$

Если система находится в равновесии, то $\Delta G_{\text{поли}} = 0$. Концентрация H^+, при которой кислота ионизована наполовину, равна

$$(H^+)_{1/2} = \exp [-(\Delta G^\circ + \Delta G_{\text{сопр}})/RT]$$

Тогда кажущаяся pK_a равна

$$pK'_a = -\lg (H^+)_{1/2} = (\Delta G^\circ + \Delta G_{\text{сопр}})/2{,}303RT$$

Если можно найти модельную систему, в которой отсутствует сопряжение процессов или взаимодействий, то для нее pK_a будет равна

$$pK_a = \Delta G^\circ/2{,}303\ RT$$

Следовательно, энергию сопряжения можно выразить через разность значений pK_a:

$$\Delta G_{\text{сопр}} = 2{,}303\ RT\ (pK'_a - pK_a)$$

Применим этот результат для анализа электростатического взаимодействия между COO^-- и NH_3^+-группами аланина, используя данные, приведенные в табл. 2.1. Величина pK_a для реакции диссоциации $COOJ \rightarrow COO^- + H^+$ определяется титрованием $(Ala)_4$; она равна 3,42, pK'_a для Ala равна 2,34. Так как при комнатной температуре $RT \approx 0{,}6$ ккал/моль, то

$$\Delta G_{\text{сопр}} = (2{,}303)(0{,}6)(2{,}34\text{-}3{,}42) = -2{,}5 \text{ ккал/моль}.$$

Таким образом, электростатическое взаимодействие между NH_3^+=и COO^--группами уменьшает на 2,5 ккал/моль свободную энергию остатка аланина. Это облегчает диссоциацию аланина

$$NH_3^+—CH(CH_3)—COOH \rightleftarrows NH_3^+—CH(CH_3)—COO^- + H^+$$

что проявляется в уменьшении pK_a.

ИОНИЗАЦИЯ БОКОВЫХ ГРУПП

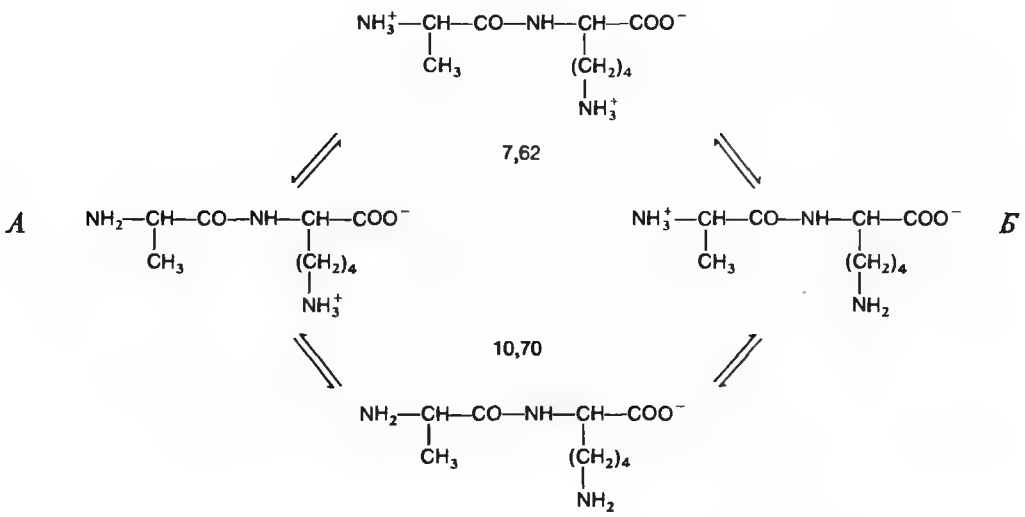
При наличии ионизуемых боковых групп данные титрования могут оказаться чрезвычайно сложными. При pH < 2 лизин существует преимущественно в форме

$$NH_3^+—CH(COOH)—(CH_2)_4—NH_3^+$$

Титрование лизина дает три значения pK_a. Одно из них, равное 2,18, соответствует (по аналогии с аланином) ионизации карбоксильной группы. По сравнению с аланином это значение pK_a смещено в сторону меньших pH вследствие более сильного притяжения между COO^-- и NH_3^+-группами молекулы. Двум другим процессам ионизации соответствуют значения pK_a, равные 8,90 и 10,28. Одно из них должно относиться к ионизации α-аминогруппы, другое — к ионизации ε-аминогруппы. Однако непосредственно определить, какая именно величина относится к данной группе, можно только путем исследования модельных соединений.

Титрование дипептида Ala — Lys дает три четкие точки ионизации, в которых значения pK_a соответственно равны 3,22, 7,62 и 10,70. Наименьшее значение pK_a соответствует депротонированию карбоксильной группы, что следует из сравнения кривых титрования двух дипептидов, Ala — Ala и Ala — Lys. Для идентификации других значений pK_a можно предложить две равновероятные модели процесса, а именно: α-аминогруппа теряет протон раньше, чем боковая цепь, и первой теряет протон боковая цепь.

Прежде чем отдать предпочтение схеме *А* или *Б* (см. с. 48), необходимо учесть следующие экспериментальные данные. Во-первых, алкилированные аминогруппы имеют более высокие pK_a, чем несущие при α-углеродном атоме карбонильную группу замещенные метиламины. Например, pK_a н-бутиламина равна 10,59, тогда как аминогруппа в α-аминокислоте, не участвующая в электростатическом взаимодействии, титруется в области pH 8 (см. данные для олигоаланинов, приведенные в табл. 2.1). Это доказывает правильность схемы *А*. Если эта схема верна, то pK_a ε-аминогруппы лизина будет несколько выше, чем pK_a бутиламиногруппы из-за ее электростатического взаимодействия с группой COO^-. И наоборот, α-NH_3^+-группа будет титроваться при наличии заряженной ε-NH_3^+-группы, и, следовательно, для нее значение pK_a должно быть ниже, чем для α-аминогруппы в дипептиде Ala — Ala. Эти предположения полностью подтверждаются экспериментальными данными. Если бы была верна схема *Б*, то она должна была бы объ-



яснить, почему α-аминогруппа в Ala — Ala титруется при pK_a 8,30, тогда как в Ala — Lys с незаряженной боковой группой α-аминогруппа титруется при pK_a 10,70. Поскольку объяснить это при помощи схемы *Б* нельзя, описываемый ею процесс не является основным при ионизации дипептида.

Для того чтобы определить, в какой точке титруется ε-аминогруппа в составе полипептида, необходимо исследовать такие соединения, у которых свободные амино- и карбоксильные концы находятся настолько далеко от боковой группы, что можно пренебречь электростатическими взаимодействиями. Например, для Ala—Lys—Ala—Ala наблюдаются три значения pK_a: 3,58, 8,01 и 10,58. Первые два очень близки к pK_a, полученным для $(Ala)_4$, а pK_a ε-аминогруппы незначительно отличается от pK_a н-бутиламиногруппы. Однако у соединений с многочисленными ионизуемыми боковыми цепями кривые титрования становятся чрезвычайно сложными. Например, Lys—Lys—Lys дает при титровании следующие значения pK_a: 3,06, 7,34, 9,80, 10,54 и 11,32.

ИОНИЗАЦИОННОЕ РАВНОВЕСИЕ В БЕЛКАХ

Приведенные выше примеры указывают на то, что титрование отдельных остатков в белке зависит от последовательности аминокислот в нем. Однако до сих пор мы имели дело лишь с короткими пептидами, обладающими относительно гибкой конформацией. Белковые же молекулы, кроме первичной, имеют вторичную и третичную структуру. Поэтому при анализе ионизационного равновесия в белках необходимо принимать во внимание электростатические взаимодействия, возникающие между всеми близко расположенными друг к другу зарядами. Попытки количественного исследования этих взаимодействий в отдельных случаях были достаточно успешными, однако в целом эта проблема до сих пор остается чрезвычайно сложной. Как будет показано ниже, все полиэлектролиты создают вокруг себя облако из противоионов, что приводит к ослаблению ионных взаимодействий. Плотность этой «ионной атмосферы» зависит от концентрации электролита в растворе и от геометрии близко расположенных заряженных групп полимера. В принципе в условиях высокой ионной силы данные титрования прежде всего отражают электростатические взаимодействия между близлежащими заряженными группами. Следовательно, при таком титровании мы получаем некоторую информацию о локальной вторичной и третичной структурах. Однако сами по себе данные полного потенциометрического титрования еще недостаточно информативны.

Кривые титрования ионов водорода типичного белка приведены на рис. 2.2. Даже для их простой интерпретации необходимо знать условия ионизационного равновесия для боковых групп всех аминокислот, входящих в состав этого белка. Как отмечалось ранее, данные потенциометрического титрования для боковых аминокислотных групп в отсутствие достаточно близко расположенных зарядов могут быть получены путем измерения pK_a пептидов. В табл. 2.2 приведены значения pK_a для разных боковых групп изолированных аминокислот; наблюдаемые для белков значения pK_a отдельных боковых групп обычно не отличаются от этих значений больше чем на единицу. В табл. 2.3 приведены данные титрования некоторых аминокислотных остатков, входящих в состав ряда белков. Из нее видно, что имеются исключения, для которых наблюдаются значительно большие изменения pK_a. Обычно это связано с гистерезисом титрования, выражающимся в том, что кривые кислотного и щелочного титрования не совпадают. Подобные эффекты возникают, когда титрование сопровождается конформационными изменениями.

Предположим, что боковые цепи непосредственно участвуют во взаимодействиях, стабилизирующих вторичную или третичную структуру. Тогда эти взаимодействия могут нарушаться при ионизации этих групп. Рассмотрим, например, остаток тирозина, связанный водородной связью с карбоксилом другого остатка. Тогда процесс титрования можно представить следующим образом:

$$OH^- + Tyr{-}O{-}H \cdots {}^-O{-}CO{-}R \rightleftarrows Tyr{-}O^- + {}^-O{-}CO{-}R + H_2O$$

Даже если в этом взаимодействии не участвует остальная часть молекулы, оно, несомненно, оказывает значительное влияние на наблюдаемое значение pK_a тирозина, так как свободная энергия водородной связи составляет, вероятно, несколько килокалорий. Если же тирозин и COO^--группа физически не могут разойтись без того, чтобы не изменялись

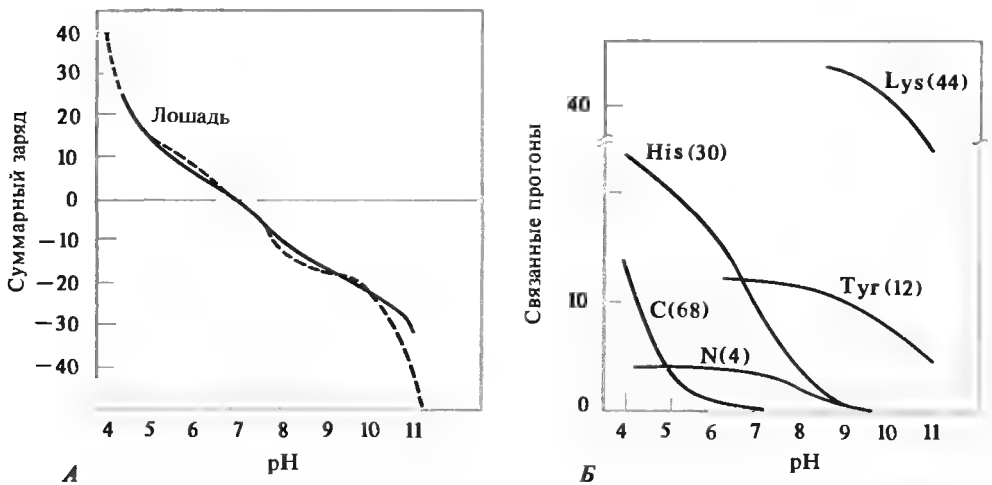
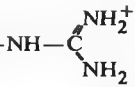
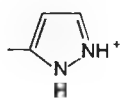
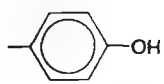


РИС. 2.2. Кривые титрования оксигемоглобина лошади при ионной силе 0,02 и температуре 29°С. *А*. Для получения правильных значений изоионных pH экспериментальные данные были увеличены на три зарядовые единицы. Теоретическая кривая учитывает изменения ионизационных характеристик всех боковых групп, вызванные электростатическими взаимодействиями в известной четвертичной структуре белка. Сплошная кривая — теория, пунктир — данные Кона и др. (1937). *Б*. Зависимость рассчитанного числа протонов, связываемых каждой ионизуемой единицей, от величины pH. С означает Glu плюс Asp, N — N-концевой участок. Числа в скобках указывают число групп данного типа. [W.R. Orttung, Biochemistry, 9, 2394 (1970).]

Таблица 2.2

ЗНАЧЕНИЯ pK_a ДЛЯ ИОНИЗУЕМЫХ ГРУПП В АМИНОКИСЛОТАХ[1)]

Аминокислота	α-COOH	α-NH_3^+	Боковая группа	pK_a в аминокислоте	Ожидаемое значение pK_a в белке
Аланин	2,3	9,9	-	-	-
Аргинин	1,8	9,0		12,5	⩾ 12
Аспарагин	2,0	8,8	-	-	-
Аспарагиновая кислота	2,0	10,0	$-COOH$	3,9	4,4-4,6
Цистеин	1,8	10,8	$-SH$	8,3	8,5-8,8
Глутаминовая кислота	2,2	9,7	$-COOH$	4,3	4,4-4,6
Глутамин	2,2	9,1	-	-	-
Глицин	2,4	9,8	-	-	-
Гистидин	1,8	9,2		6,0	6,5-7,0
Изолейцин	2,4	9,7	-	-	-
Лейцин	2,4	9,6	-	-	-
Лизин	2,2	9,2	$-NH_3^+$	10,8	10,0-10,2
Метионин	2,3	9,2	-	-	-
Фенилаланин	1,8	9,1	-	-	-
Пролин	2,0	10,6	-	-	-
Серин	2,1	9,2	-	-	-
Треонин	2,6	10,4	-	-	-
Триптофан	2,4	9,4	-	-	-
Тирозин	2,2	9,1		10,9	9,6-10,0
Валин	2,3	9,6	-	-	-

[1)] Значения pK_a в большинстве случаев получены при 25°C. Ожидаемое pK_a в белках определяется при помощи модельных соединений, в которых титрование боковых цепей не зависит от влияния α-заместителей.

также положения других остатков, то для ионизации потребуются дополнительные энергетические затраты, обусловленные либо отталкиванием отрицательно заряженных и близко расположенных групп Tyr—O^- и ^-O—CO—R, либо изменением конформации других остатков. Полная свободная энергия ионизации равна

$$\Delta G_{\text{полн}} = \Delta G_{\text{иониз}} + \Delta G_{\text{сопр}} \qquad (2.1)$$

где $\Delta G_{\text{иониз}}$ — свободная энергия ионизации изолированного тирозина, а $\Delta G_{\text{сопр}}$ — свободная энергия, обусловленная разрывом водородной связи, электростатическим отталкиванием и конформационными изменениями. Как видно из Дополнения 2.1, влияние $\Delta G_{\text{сопр}}$ заключается в изменении pK'_a. Поскольку $\Delta G_{\text{сопр}} > 0$, pK'_a смещается до бо́льших значений, чем ожидаемые в случае титрования изолированного тирозина.

Таблица 2.3

pK_a, НАБЛЮДАЕМЫЕ ДЛЯ ИОНИЗУЕМЫХ ГРУПП В НЕКОТОРЫХ БЕЛКАХ[1)]

Аминокислота или группа	Белок	pK_a[2)]
Аспарагиновая и глутаминовая кислоты	Лизоцим (3 из 16 групп)	Д
	β-Лактоглобулин (49 из 51 группы)	4,8
	(2 из 51 группы)	7,3
	Сывороточный альбумин	4,0
	Инсулин	4,7
Гистидин	Миоглобин (6 из 12 групп)	Д
	(6 из 12 групп)	6,6
	β-Лактоглобулин	7,4
	Лизоцим	6,8
	Инсулин	6,4
	Сывороточный альбумин	6,9
Тирозин	Химотрипсиноген (1 из 4 групп)	9,7
	(1 из 4 групп)	10,4
	(2 из 4 групп)	П
	Инсулин	9,6
	Сывороточный альбумин	10,4
	Рибонуклеаза (3 из 6 групп)	9,6
Лизин	Сывороточный альбумин	9,8
	Лизоцим	10,4
	Химотрипсиноген (3 из 13 групп)	П
Аргинин	Инсулин	11,9
α-Карбоксил	-"-	3,6

[1)] R. H. Haschemeyer, A. E. V. Haschemeyer, Proteins (New York: Wiley, 1974).
[2)] Д — депротонирована при всех значениях pH, П — протонирована при всех значениях pH.

Если конформационное изменение, сопровождающее процесс ионизации, обратимо, то такой же эффект наблюдается и при проведении обратного титрования. Часто, однако, возвращение к нативной конформации происходит очень медленно или даже оказывается невозможным. В таких случаях при обратном титровании остатки тирозина титруются нормально, так как отсутствует связь между процессом ионизации и конформационными изменениями. Более детальные экспериментальные данные приведены в гл. 7, где описано титрование остатков тирозина, входящих в состав рибонуклеазы, причем в данном случае специфическая ионизация этих остатков может наблюдаться оптическими методами.

Для того чтобы наилучшим образом проанализировать кривую титрования (см., например, рис. 2.2), нужно сопоставить ее с известными данными по аминокислотному составу. Различные участки кривой ставят в соответствие известному числу способных к ионизации остатков (взятому с определенным весом) в предположении, что последние титруются нормально. Если при этом не удается получить соответствия между наблюдаемым и ожидаемым числом присоединенных или потерянных протонов, то делается предположение, что некоторые остатки не титруются совсем, поскольку они участвуют в стабилизации конформации белка. Для более тонкого анализа необходимо учитывать сдвиг pK_a из-за электростатического взаимодействия. Достоверные значения такого сдвига можно получить только в том случае, если известна третичная структура. Как видно из рис. 2.2, при благоприятной ситуации характер кривой титрования можно довольно хорошо объяснить количественно.

ПОЛЯРНОСТЬ БОКОВЫХ ГРУПП АМИНОКИСЛОТ

Из всего сказанного ясно, что можно четко разграничить заряженные и незаряженные аминокислоты при любых значениях pH. Было бы полезно располагать и другими способами однозначной классификации аминокислот, основанными на различиях в их свойствах. Можно попытаться сделать это, исходя из интуитивных химических представлений о полярности боковых групп. Аминокислоты с неполярными боковыми группами хуже растворяются в воде. В этот класс, очевидно, входят аминокислоты с алифатическими углеводородными боковыми цепями: аланин, валин, лейцин и изолейцин. К неполярным аминокислотам можно отнести также финилаланин, триптофан и метионин.

На другом конце шкалы полярности находятся Glu, Asp, Arg и Lys. Полярны также аспарагиновая и глутаминовая кислоты, содержащие амидную боковую группу. Гидроксилсодержащие аминокислоты — серин и треонин — также полярны из-за большого дипольного момента и способности гидроксила образовывать водородные связи. Все эти молекулы активно взаимодействуют с водой и хорошо в ней растворяются.

Полярность оставшихся пяти аминокислот не столь ясна. Значения pK_a для цистеина и гистидина настолько близки к 7, что во многих белках при физиологических условиях эти аминокислоты действительно могут быть заряженными. Таким образом, они могут быть полярными, но если они незаряжены, то они совершенно теряют свои полярные свойства. В тирозине ароматическое кольцо компенсирует влияние гидроксильной группы и обычно Tyr ведет себя как неполярный остаток. Глицин и пролин — особые аминокислоты. Их структура значительно отличается от структуры остальных 18 остатков. Глицин лишен оптически активного асимметричного α-углерода, откуда следует, что он может существовать в конформационных состояниях, недоступных другим аминокислотам. Он не имеет боковой группы, но его поведение в составе белков позволяет отнести его к категории неполярных остатков. Пролин же на самом деле является имино-, а не аминокислотой. Поэтому его конформационные свойства отличаются от свойств всех остальных остатков. Трудно что-нибудь сказать о его полярности, но, учитывая его локализацию в белках, можно прийти к заключению, что Pro чаще ведет себя как полярный остаток.

В табл. 2.4 приведены данные по растворимости различных аминокислот в воде при нейтральных pH, которые дают возможность проверить наши полуинтуитивные рассуждения. Используя принцип «подобное растворяется в подобном», а также то обстоятельство, что молекулы воды полярны, можно сказать, что большая часть представленных данных качественно согласуется с нашими выводами. Однако к этим данным не следует подходить с точными количественными мерками, так как в табл. 2.4 приводятся растворимости цвиттерионных форм аминокислот, тогда как нас интересуют свойства аминокислотных боковых групп как составных частей пептида.

Существуют и более серьезные ограничения при использовании данных по растворимости аминокислот. Растворение твердого вещества в воде — равновесный процесс. На него влияют и сродство этого вещества к воде, и сродство молекул друг к другу в кристалле. Нас интересуют здесь только взаимодействия первого типа, но мы не уверены, что взаимодействия второго типа не играют важной роли в процессе растворения. В принципе можно было бы обратиться к другим данным, чем те, которые приведены в табл. 2.4, например, к данным о растворимости газообразных аминокислот или их боковых групп в воде. Получить такую информацию с помощью прямых измерений невозможно, однако есть другой, косвенный подход. Так, зная термодинамику испарения данного вещества с поверхности кристалла и термодинамику растворения твердого вещества, можно получить представление об интересующем нас равновесии между паром и раствором.

Тэнфорд предложил альтернативный подход, который в отсутствие данных о газообразной фазе можно использовать для простого термодинамического описания свойств

Таблица 2.4

КЛАССИФИКАЦИЯ СВОЙСТВ АМИНОКИСЛОТ

Остаток	Тип аминокислоты	Растворимость цвиттерионной формы при 25°С, моль/кг	$\Delta G_{пер}$, ккал/моль EtOH → H_2O
Trp	Неполяриая	0,07	3,00
Ile	Неполярная	0,26	2,95
Tyr	Неполярная [3]	< 0,00	2,85
Phe	Неполярная	0,17	2,65
Leu	Неполярная	0,16	2,40
Val	Неполярная	0,50	1,70
Met	Неполярная	0,38	1,30
Cys [1]	Неполярная	< 0,00	1,00
Ala	Неполярная	1,86	0,75
Gly	Неполярная [3]	3,33	0,00
His	Не определен	—	—
Pro	Полярная [3]	14,1	2,60
Ser	Полярная	4,02	—
Thr	Полярная	—	0,45
Asn	Полярная	0,19	—
Glu	Полярная	0,29	—
Asp	Заряженная	0,04 [2]	—
Glu	Заряженная	0,06 [2]	—
Lys	Заряженная	3,95 [4]	1,50
Arg	Заряженная	4,06 [4]	0,75

[1] Приводятся данные для полуцистина.
[2] При наличии протонированной боковой цепи; в отсутствие протона эта величина будет намного выше.
[3] Обычно.
[4] Соли HCl.

того окружения, которое предпочитают разные аминокислоты. Он сравнивал растворимость аминокислот в воде (как в типичном полярном растворителе) и в этаноле. Этанол служил моделью неполярного растворителя, все еще обладающего способностью растворять аминокислоты в количестве, достаточном для проведения измерений. Так как в обоих растворах взаимодействия с кристаллической фазой одинаковы, можно использовать данные о свободной энергии растворения при расчете ΔG для перехода 1 моль остатков из этанола в воду при постоянной концентрации. ΔG, естественно, будет зависеть от взаимодействия растворителя как с боковыми цепями, так и с заряженными амино- и карбоксильными группами. В качестве стандартной аминокислоты, лишенной боковой группы, Тэнфорд избрал глицин. Вклад боковой группы в свободную энергию переноса Тэнфорд определил как разность между ΔG для соответствующей аминокислоты и ΔG для глицина. Для этого имеются разумные основания, поскольку с достаточной степенью точности можно считать свободную энергию переноса аддитивной функцией структуры молекулы.

В гл. 5 мы рассмотрим вопрос о свободной энергии переноса значительно подробнее. Здесь же мы приведем только некоторые результаты, необходимые для понимания проблемы в целом. Растворимость аминокислот в этаноле значительно меньше, чем в воде. Для глицина $\Delta G = -4{,}63$ ккал/моль; отсюда следует, что если бы оказалось возможным

проведение прямого эксперимента, то глицин самопроизвольно перешел бы из этанола в воду. Для фенилаланина $\Delta G = -1{,}98$ ккал/моль. Разность между этими двумя величинами, равная $+2{,}65$ ккал/моль, представляет свободную энергию переноса $\Delta G_{\text{пер}}$ боковой группы фенилаланина. Отметим, что она положительна, так же как и ΔG боковых групп большинства других аминокислот (см. табл. 2.4). Таким образом, изолированная боковая цепь, не «отягощенная» остальной частью аминокислоты, могла бы самопроизвольно переходить из воды в этанол. Ясно, что подобные боковые группы будут предпочтительно располагаться во внутренней области свернутого белка или по крайней мере избегать водного окружения.

Значения $\Delta G_{\text{пер}}$ отражают действие гидрофобных сил, играющих важную роль в определении третичной структуры белков. В гл. 5 подробно обсуждается все, что в настоящее время известно о природе этих сил. Данные о $\Delta G_{\text{пер}}$ можно использовать для разделения аминокислот на гидрофобный и гидрофильный классы. Как видно из табл. 2.4, представления об этих категориях неплохо согласуются с простыми интуитивными представлениями о полярности аминокислот. Казалось бы, этанол выбран в качестве эталонного неполярного растворителя произвольно, однако выбор другого растворителя лишь в незначительной степени изменяет основные выводы.

2.2. Химический состав белков

Белки состоят из 20 обычных аминокислот. Иногда при выделении и анализе белка мы обнаруживаем только эти аминокислоты. Однако зачастую в белках находят и другие соединения, или ковалентно связанные с белковой молекулой, или присоединенные к ней другим образом, но так прочно, что их невозможно отделить при очистке. В этом разделе мы попытаемся суммировать все, что касается состава типичных белков.

АМИНОКИСЛОТНЫЙ СОСТАВ

Прежде всего рассмотрим сами аминокислоты. Автоматические методы анализа аминокислотного состава белков в настоящее время доведены до высокого совершенства, в результате чего мы располагаем огромным фактическим материалом. Несколько основных особенностей аминокислотного состава белковых молекул было выявлено в самом начале исследований. Так, встречаемость данной аминокислоты сильно меняется от белка к белку, однако в общем некоторые аминокислоты (такие, как метионин и триптофан) встречаются относительно редко. Другие (аланин и лейцин), как правило, встречаются чаще всего. На рис. 2.3 представлен аминокислотный состав ряда белков с совершенно разными биологическими свойствами.

Гистоны представляют собой небольшие основные белки, связывающиеся с ДНК в хроматине. Они содержат относительно много положительно заряженных аминокислот, которые обеспечивают известную электростатическую стабилизацию хроматина путем взаимодействия с отрицательно заряженными фосфатами ДНК.

Белок пурпурных мембран in vivo связан с мембранами и, как полагают, в основном локализуется на внутренней стороне липидного бислоя. Поэтому вполне логичным представляется тот факт, что он содержит относительно мало заряженных аминокислот, но зато весьма богат неполярными остатками.

При сравнении состава других белков, приведенных на рис. 2.3, видно, что для небольших глобулярных или фибриллярных белков (в отличие от крупных глобулярных) характерна тенденция к увеличению содержания полярных аминокислот. Это можно объяснить, исходя из отношения площади поверхности к объему молекулы. Чем больше это отношение, тем больше должно быть число полярных остатков, поскольку, как мы увидим ниже, они стремятся располагаться в основном у поверхности белковой глобулы. Предло-

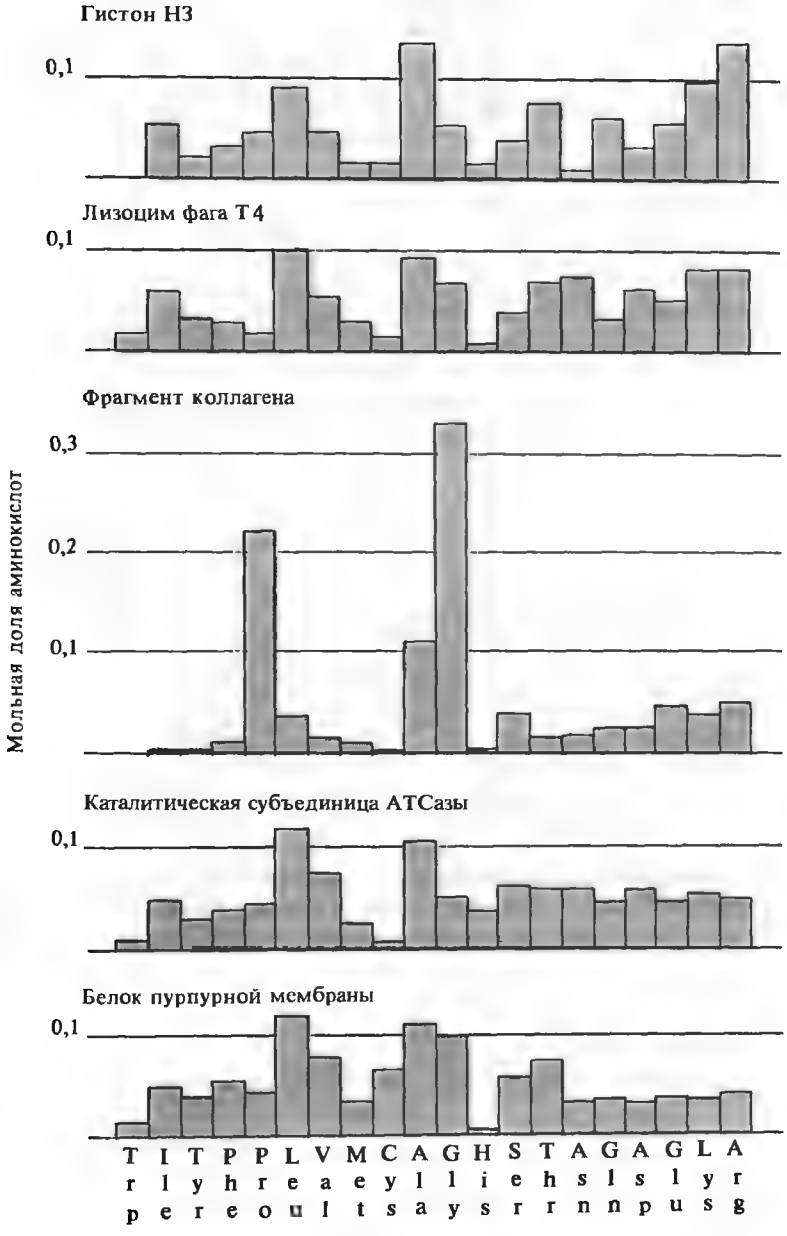


РИС. 2.3. Аминокислотный состав пяти белков. Порядок расположения аминокислот примерно соответствует убыванию их гидрофобности.

жить какую-либо содержательную концепцию относительно среднего аминокислотного состава белков довольно трудно. Можно лишь отметить, что аминокислотный состав белков данного организма или данной ткани находится посередине между экстремальными значениями, характерными для сильно полярных и совершенно неполярных белков. Приведенные на рис. 2.3 примеры специально подобраны так, чтобы подчеркнуть наблюдаемые различия в составе белков. В пределах каждого данного класса белков эти различия настолько велики, что затрудняют простое их разбиение на классы полярных и неполярных.

Определение аминокислотного состава белков — большая область исследований, в которой данные относительно легко получить, но нельзя однозначным образом использовать для извлечения полезной информации о свойствах структуры. Были предприняты многочисленные попытки рассчитать средний аминокислотный состав отдельных белков в надежде, что подобные средние в какой-то степени коррелируют с основными свойствами этих макромолекул. Иногда исследователь располагает лишь данными об аминокислотном составе. Поэтому представляется весьма желательным научиться использовать эти данные хотя бы для планирования дальнейших исследований.

Поскольку величину $\Delta G_{\text{пер}}$ для данной аминокислоты можно использовать в качестве меры гидрофобности (см. табл. 2.4), мы можем построить шкалу средней гидрофобности H_ϕ:

$$\bar{H}_\phi = \sum_i \Delta G^i_{\text{пер}} \chi_i \tag{2.2}$$

где $\Delta G^i_{\text{пер}}$ — свободная энергия переноса для аминокислотного остатка i-го типа, а χ_i — мольная доля последнего в данном белке. Величина $\bar{H}_\phi$, вычисленная для 150 исследованных белков и пептидов, колеблется в пределах от 0,44 до 2,02 ккал · (моль остатков)$^{-1}$. Однако, как показывают данные рис. 2.4, значения средней гидрофобности большинства белков группируются в узкой области близ 1,00 ккал · (моль остатков)$^{-1}$. Следовательно, подобный параметр не может принести большую пользу при проведении четкого различия между молекулами белков.

Значение величины $\bar{H}_\phi$ определяется главным образом мольной долей некоторых из неполярных боковых групп. На основе простого расчета отношения встречаемостей χ каких-либо представляющих интерес боковых групп было построено несколько других шкал. Например, в выражении $R = \sum_k \chi_k / \sum_j \chi_j$ k может соответствовать гидрофильным, а j — гидрофобным боковым группам либо k — полярным, а j — неполярным и т.д. Одна из особенностей белковых структур состоит в довольно устойчивой тенденции заряженных или сильно полярных остатков располагаться у поверхности молекул. Если счи-

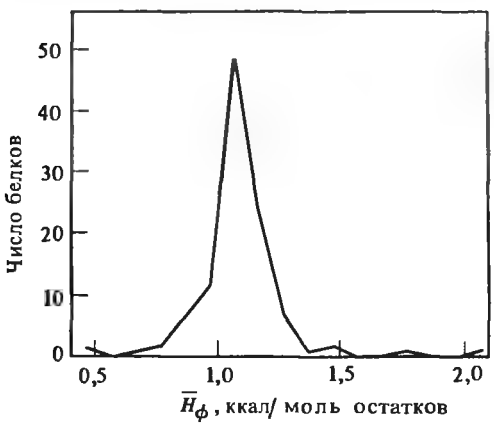


РИС. 2.4. Средняя гидрофобность $\bar{H}_\phi$ рассчитанная для 150 белков в соответствии с уравнением (2.2). [C.C. Bigelow, J. Theoret. Biol., **16**, 187, (1967).]

тать,что k соответствует Arg, Lys, His, Gly, Glu, Asp, Asn, а j — Ile, Tyr, Phe, Leu, Val и Met, можно построить шкалу R_3, в которой будет в сильной степени отражена эта тенденция. Конкретный выбор остатков для расчета величины R_3 до некоторой степени произволен. Однако это отношение может служить весьма полезным критерием, так как значения R_3 для разных белков существенно различаются. Так, R_3 для 13 белков изменялся от 0,36 до 2,03, тогда как параметр H_ϕ для тех же белков лежал в пределах от 0,86 до 1,29 ккал · моль$^{-1}$. Хотя между величинами R_3 и $\bar{H}_\phi$ имеется некоторая антикорреляция, нельзя сказать, что они меняются в противоположных направлениях. Следовательно, используя одновременно оба показателя, мы получаем больше информации об исследуемой системе. Способ, каким лучше использовать эти шкалы, зависит от конкретной задачи, а чтобы найти оптимальную комбинацию параметров R_3 и H_ϕ, нужно использовать статистические методы.

ПРЕДСКАЗАНИЕ СВОЙСТВ БЕЛКОВ ПО АМИНОКИСЛОТНОМУ СОСТАВУ

Анализируя данные по аминокислотному составу мембранных белков, можно установить, является ли тот или иной белок внутренним (т.е. содержащим большое количество неполярных и относительно мало заряженных остатков) или внешним. В последнем случае мембранный белок весьма похож на типичный растворимый немембранный белок. Не так давно Ф. Баррантес сравнил величины R_3 для большого числа мембранных белков с соответствующими значениями H_ϕ. При расчете $\bar{H}_\phi$ он исключил данные для триптофана. Из результатов его исследований, приведенных в табл. 2.5, видно, что R_3 для внутренних и внешних мембранных белков сильно отличаются друг от друга, чего нельзя сказать о величинах H_ϕ.

Можно задаться вопросом, какая линейная комбинация R_3 и $\bar{H}_\phi$ оптимальна для разграничения внутренних и внешних мембранных белков? Соответствующая функция Z называется дискриминантной и записывается в виде

$$Z = -0{,}345R_3 + 0{,}60H_\phi. \qquad (2.3)$$

Средние значения Z для внутренних и внешних мембранных белков равны соответственно 0,52 ± 0,11 и 0,12 ± 0,16. Для немембранных белков среднее значение Z составляет 0,16 ± ± 0,17. Если величина Z некоторого белка находится в пределах указанных стандартных отклонений, то вероятность ошибки в определении его характера не превышает 8%.

В качестве примера использования параметра Z для выявления некоторых особенностей структурной организации белков рассмотрим результаты, приведенные в табл. 2.6. Величины H_ϕ и Z рассчитаны для нескольких мембранных белков, включая два, относи-

Таблица 2.5

СРАВНЕНИЕ R_3 И H_ϕ ДЛЯ НЕКОТОРЫХ БЕЛКОВ[1)]

Средняя переменная	205 немембранных белков	24 мембранных белка	
		считающиеся внутренними	считающиеся внешними
R_3	1,26 ± 0,42	0,59 ± 0,18	1,37 ± 0,35
$\bar{H}_\phi$	0,996 ± 0,098	1,197 ± 0,097	0,986 ± 0,075

[1)] F. J. Barrantes, Biochem. Biophys. Res. Commun., **62**, 407 (1975).

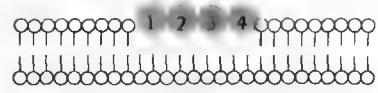
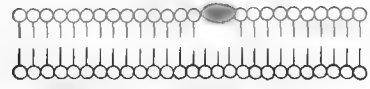
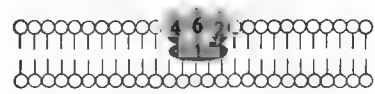
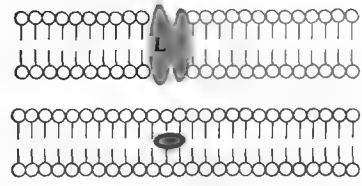
тельно которых известно, что они состоят из нескольких субъединиц. Прежде всего интересно знать, где скорее всего расположены отдельные субъединицы: внутри или снаружи мембран? Все субъединицы белка рецептора ацетилхолина и четыре субъединицы цитохромоксидазы (для которых есть соответствующие данные) имеют сходные значения H_ϕ. Однако значения параметра Z говорят о том, что одна субъединица цитохромоксидазы должна располагаться внутри, а три другие — на поверхности мембраны (см. табл. 2.6). Значения Z для субъединиц белка рецептора ацетилхолина не отличаются друг от друга. Отсюда следует, что все они локализованы в мембране в одном и том же месте. Утверждение, что параметр Z является строгим структурных критерием, было бы слишком смелым, но как грубый показатель организации субъединиц в сложной четвертичной структуре он, видимо, вполне пригоден.

МНОГИЕ БЕЛКИ СОДЕРЖАТ НЕОБЫЧНЫЕ АМИНОКИСЛОТЫ

Если мы хотим выяснить некоторые особенности структуры и функции белков исходя из их химического состава, то нужно рассматривать не только аминокислоты, но и другие соединения, которые могут входить в их состав. Молекула белка, после того как она синтезировалась в рибосоме, может претерпевать различные изменения. В самом простом случае происходит разрыв одной или нескольких пептидных связей, что приводит к обра-

Таблица 2.6

АМИНОКИСЛОТНЫЙ СОСТАВ НЕКОТОРЫХ МЕМБРАННЫХ БЕЛКОВ[1)]

Белок	H_ϕ, ккал	Z	Схематическое представление
Рецептор ацетилхолина	1,184	0,37	
Субъединица 1	1,182	0,38	
Субъединица 2	1,119	0,29	
Субъединица 3	1,140	0,31	
Субъединица 4	1,140	0,31	
Ацетилхолинэстераза	1,059	0,25	
Цитохромоксидаза	1,185	0,45	
Субъединица 1	1,185	0,51	
Субъединица 2	1,011	0,25	
Субъединица 4	1,033	0,16	
Субъединица 6	1,029	0,14	
Na^+ — K^+-АТРаза	1,272	0,42	
Большая субъединица	1,248	0,42	
Малая субъединица	1,309	0,44	
Родопсин (бычий)	1,208	0,51	
Белок пурпурной мембраны	1,247	0,56	

[1)] H_ϕ — средняя гидрофобность, определенная уравнением (2.2). Z — дискриминантная функция, определенная уравнением (2.3), (линейная комбинация средней гидрофобности и отношения числа заряженных аминокислот к числу гидрофобных). [M. A. Raftery, R. L. Vandlen, K. L. Reed, T. Lee, Cold Spring Harbor Symp. Quant. Biol., **40**, 193 (1976).]

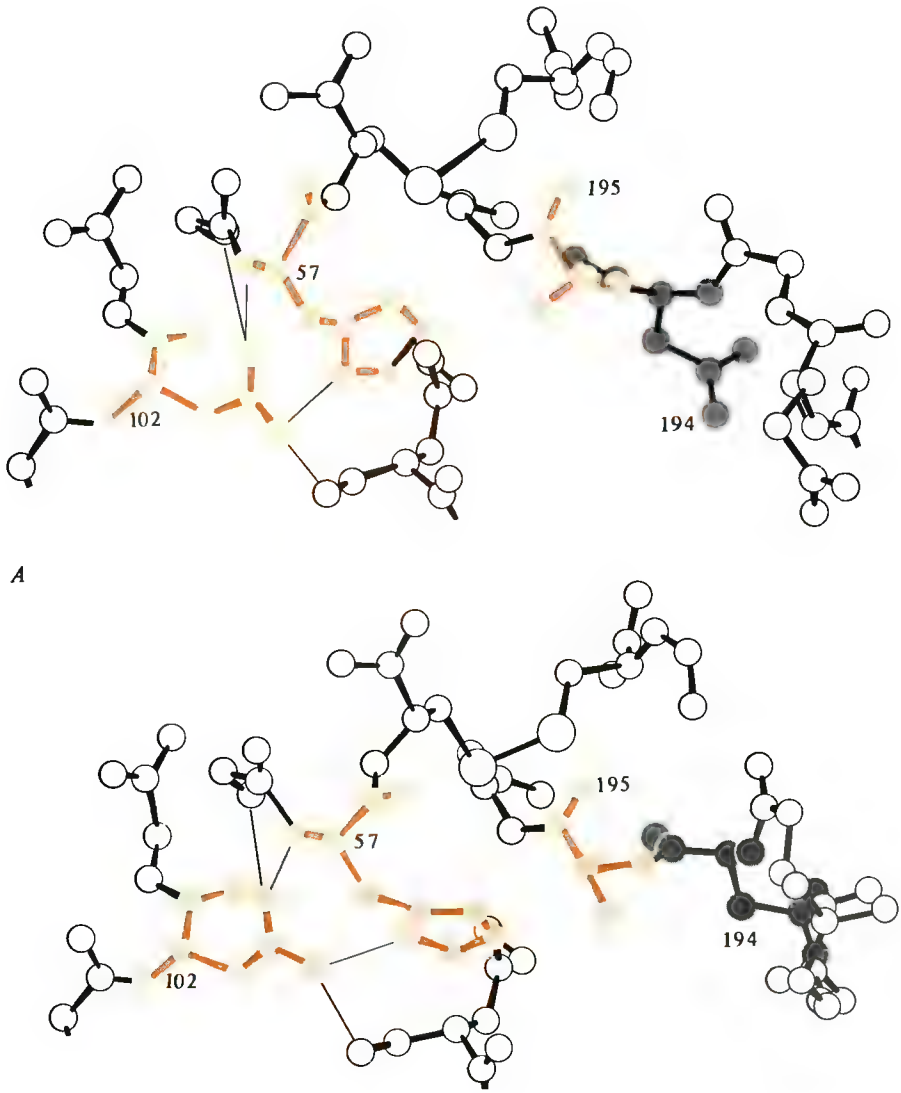


РИС. 2.5. Различие в структуре области активного центра трипсиногена (*А*) и трипсина (*Б*). В катализе непосредственно участвуют аминокислотные остатки His^{57}, Asp^{102} и Ser^{195}. При расщеплении пептидной цепи трипсиногена с образованием трипсина связь Lys^{15}-Ile^{16} гидролизуется и появляется заряженный N-конец на Ile^{16}. Этот остаток меняет свое положение и образует солевой мостик с Asp^{194}, в результате чего меняется взаимная ориентация каталитически активных остатков. Кроме того, значительные сдвиги претерпевают остатки, расположенные с правой стороны от активного центра, что в свою очередь открывает область связывания белка для субстрата. [A.A. Kossiakoff et al., Biochemistry, **16**, 654, (1977).]

зованию многоцепочечного белка. Этот процесс способен резко менять структуру и функцию белковой молекулы. В экстремальных случаях, когда отдельные полипептидные цепи связаны только нековалентными связями, эти цепи могут диссоциировать. Однако гораздо чаще происходят менее существенные изменения структуры. Если расщепление пептидной связи происходит при pH 7, то в белковой молекуле появляются концевые заряженные NH_3^+- и COO^--группы, которые сильно отличаются друг от друга по своим свойствам и, следовательно, по-разному взаимодействуют со своим окружением. Рассмотрим, например, белок трипсиноген — предшественник трипсина. Разрыв только одной пептидной связи вызывает появление чрезвычайно высокой ферментативной активности. Исходя из трехмерной структуры молекул данного фермента и его предшественника, можно видеть, что разрыв цепи позволяет аминокислотным остаткам, локализованным вблизи активного центра, существенно изменить свое местоположение. На рис. 2.5 показано, как перемещается образовавшаяся при разрыве цепи NH_3^+-группа. Кроме того, изменяет свое положение и ряд других остатков. В результате всех этих структурных изменений открывается активный центр, что позволяет молекуле фермента связываться с субстратом.

При другом посттрансляционном процессе изменяются сами аминокислоты. Несколько таких примеров показано на рис. 2.6. В ряде случаев функциональное и структурное значение этих процессов неизвестно. Однако такое изменение, как образование заряженной группы у серина в результате его фосфорилирования, может существенно повлиять на локальную структуру белка. Некоторые из этих модификаций представляют только первый этап в более сложных перестройках белковой молекулы. Например, многие белки содержат ковалентно связанные углеводные остатки, начиная от одной-двух молекул сахара, локализованных в одном месте, и кончая множеством крупных олигосахаридных включений. Некоторые из известных способов присоединения углеводов изображены на рис. 2.6. Наиболее предпочтительными местами таких присоединений являются оксиаминокислоты. Структурные последствия присоединения углеводов и формирования таким образом гликопротеидов еще недостаточно ясны. Олигосахариды чрезвычайно гидрофильны, что может оказывать известное влияние на свойства комплекса. Многие мембранные белки содержат значительное количество связанных сахаров. Быть может, углеводы, имеющие сильное сродство к водной фазе, обеспечивают «вытягивание» углеводсодержащей части белка из липидного бислоя.

Углеводные структуры значительно жестче пептидных, но до сих пор неясно, играет ли их жесткость какую-либо роль в функционировании белка. Известно, что по крайней мере некоторые гликопротеиды устойчивы к необратимой тепловой денатурации, однако роль сахаров в этом пока также неясна. И все-таки мы располагаем кое-какими данными о функциональной роли остатков сахара в белках. Именно они в первую очередь определяют некоторые иммунологические свойства. Они же, вероятно, участвуют в механизме межклеточного узнавания и связаны с многочисленными явлениями, происходящими на поверхности клетки. Углеводные части немембранных белков, видимо, являются факторами, определяющими секрецию белков из клетки, их появление в сыворотке (либо в иной внешней среде) и окончательное удаление из системы в результате процесса деградации.

МЕТАЛЛЫ И ПРОСТЕТИЧЕСКИЕ ГРУППЫ

Некоторые белки содержат простетические группы неаминокислотной природы, определенно имеющие значение для структуры белков и их функции. Например, имеется большой класс белков, для биологической активности которых необходимо наличие одного или более ионов металла. Видимо, многим белкам необходимо присутствие в их составе нековалентно связанного с белком одно- либо двухвалентного ионов щелочных металлов. Ca^{2+}-связывающий участок одного из таких белков показан на рис. 2.7, *А*. Эти ионы при-

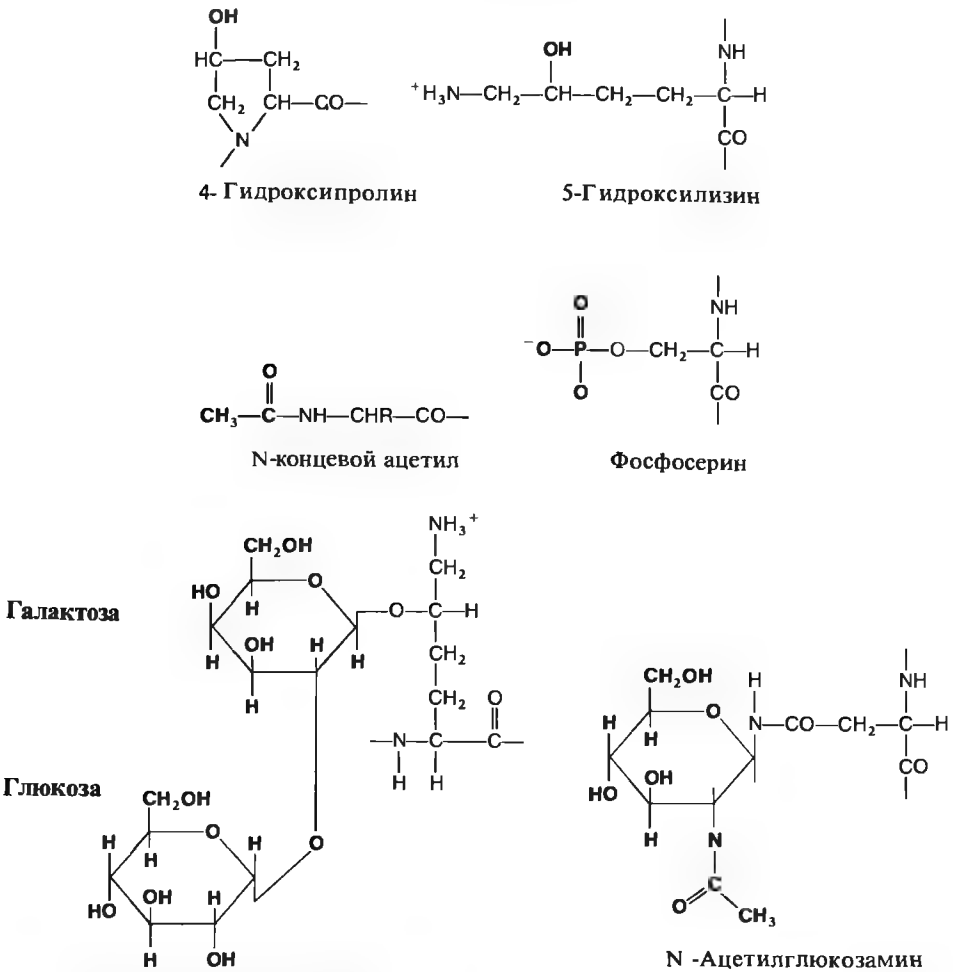


РИС. 2.6. Структура некоторых модифицированных боковых групп аминокислот. Все эти изменения происходят после того, как соответствующие обычные аминокислоты встраиваются в пептидную цепь в рибосоме.

сутствуют в клетке в больших количествах, белки же как бы приспосабливаются к таким условиям и извлекают из них максимальную выгоду. По-видимому, в большинстве случаев эти ионы принимают участие только в стабилизации структуры белка.

Сложнее обстоит дело с переходными металлами. В виде свободных катионов они встречаются в биологических системах очень редко, что важно для клетки, так как многие белки необратимо разрушаются при контакте с катионами тяжелых металлов. Белкам непросто было приспособиться к использованию преимуществ образования комплексов с такими катионами, как Fe^{2+}, Zn^{2+} и Cu^{2+}. Ими используются два основных приема. Некоторые белковые молекулы содержат участки, состоящие только из таких аминокислот, которые допускают образование прочных специфических связей с отдельными катионами. В качестве примера укажем Zn^{2+}- или Cu^{2+}-связывающие участки в ряде ферментов. На рис. 2.7, *Б* изображена структура одного из таких комплексов.

В других белках ионы металла образуют комплекс с одной или несколькими группами лиганда, ассоциированного с белком. Иногда лиганд как бы изолирует атом металла. Наиболее изученный пример такого рода — гемсодержащие белки. В таких белках содержится четырехвалентный лиганд — порфириновое кольцо. Комплекс порфирина с некоторыми металлами имеет два дополнительных места связывания, выше и ниже плоскости порфиринового кольца. В гемоглобине и миоглобине одно из этих мест занимает гистидин самого белка, тогда как другое остается свободным и может быть занято таким лигандом, как кислород (рис. 2.7,*Г*). Более сложные структуры содержатся в белках, связанных с ионами разных металлов.

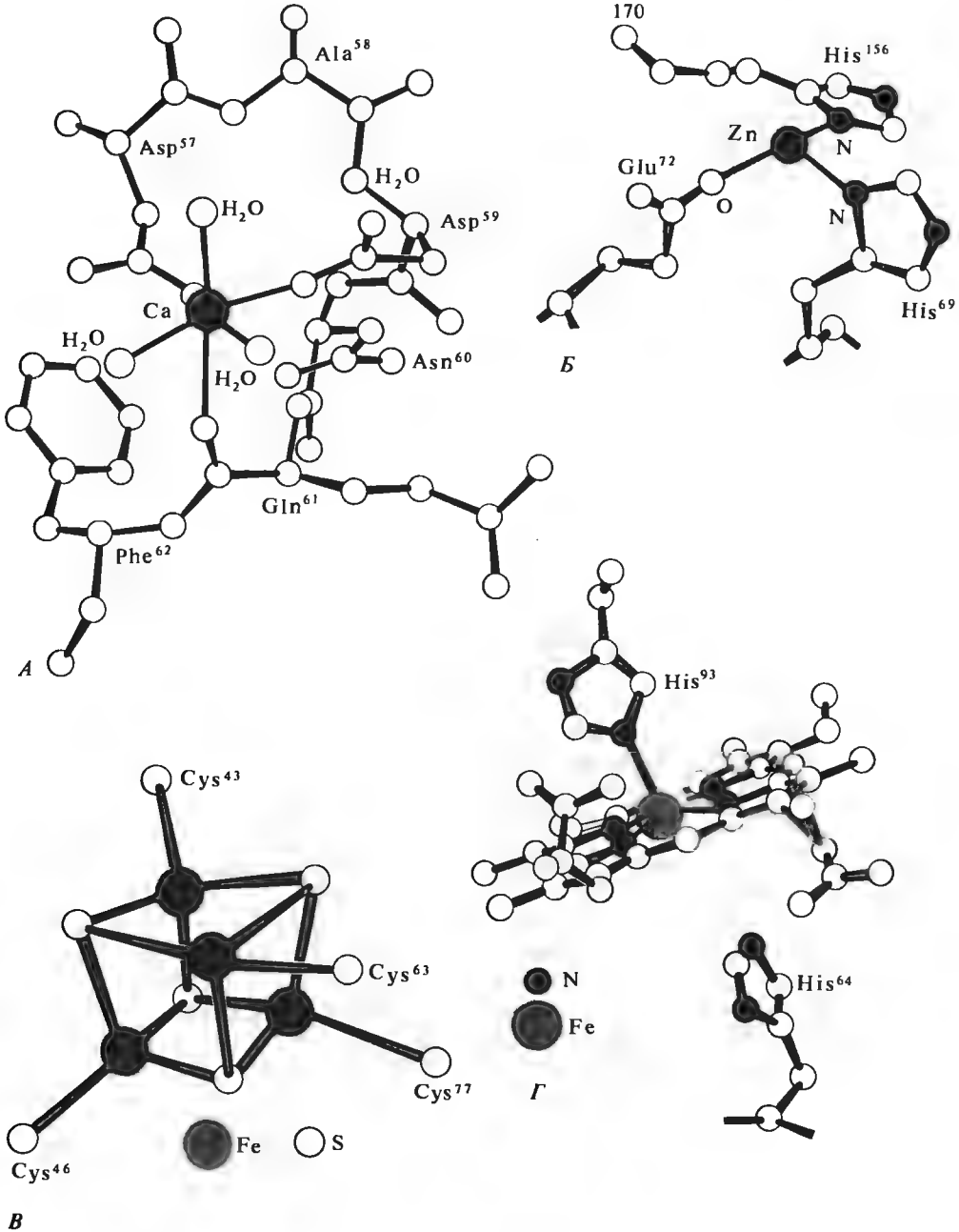
Одним из путей разрешения вопроса о том, насколько сильно влияет металл (и связанные с ним лиганды) на структуру белковой молекулы, может служить сравнение холофермента (содержащего металл) с соответствующим апоферментом (тем же белком, но без простетической группы). У гемсодержащих белков, таких, как гемоглобин и миоглобин, эти структуры очень сильно отличаются друг от друга, что не удивительно, поскольку в нативной молекуле с гемом контактирует большое число аминокислотных остатков. Удаление гема оказывает слишком сильное влияние на окружающую его часть молекулы, чтобы структура в целом могла остаться неизменной.

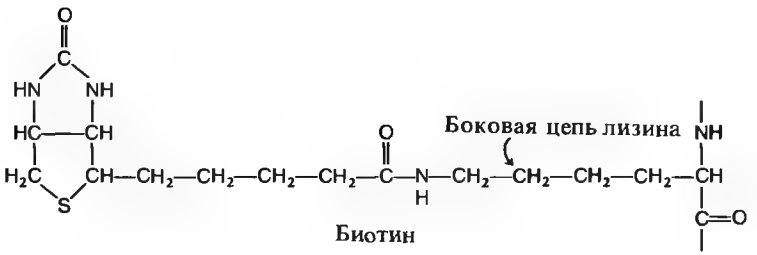
Г. Бейхок сравнил ряд свойств гемсодержащих субъединиц гемоглобина и соответствующих глобиновых полипептидных цепей. Характер ассоциации субъединиц оказался существенно различным. Ниже приведены некоторые равновесные состояния субъединиц гемоглобина. Индекс h относится к гемсодержащим цепям, индекс o — к цепям без гема. Общая картина достаточно сложна. Например, наличие гема способствует ассоциации β-субъединиц и ингибирует ассоциацию α-субъединиц. Более того, очевидно, что гем может влиять на образование четвертичной структуры, даже если он и не участвует непосредственно в этом процессе в качестве компонента, входящего в состав реакционноспособной области субъединиц. Его действие, по-видимому, сводится в основном к тому, что он изменяет вторичную и третичную структуры субъединиц.

Гемоглобин	$2\alpha_h\beta_h \rightleftarrows (\alpha_h\beta_h)_2$
Изолированные β-субъединицы	$2(\beta_h)_2 \rightleftarrows (\beta_h)_4 \qquad 2(\beta_o)_2 \rightleftarrows (\beta_o)_4$
Изолированные α-субъединицы	$2(\alpha_h) \rightleftarrows (\alpha_h)_2 \qquad 8\alpha_o \rightleftarrows 4(\alpha_o)_2 \rightleftarrows (\alpha_o)_8$
Реассоциация при 4°C	$\alpha_o + \beta_o \rightarrow \alpha_o\beta_o$
	$2\alpha_o + 2\beta_h \rightarrow (\alpha_o)_2(\beta_h)_2$
	$\alpha_h + \beta_o \rightarrow \alpha_h\beta_o$

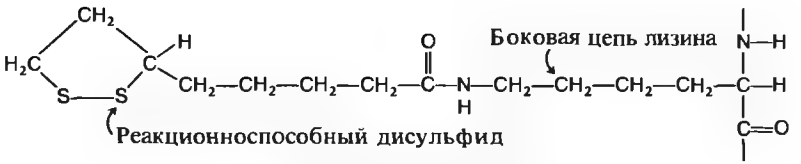
В отличие от гемоглобина многие другие белки, просто связывающие изолированный катион, при удалении из них последнего не претерпевает значительных структурных изме-

РИС. 2.7. Некоторые типичные металлсвязывающие участки в белках. На этом рисунке показаны реальные трехмерные структуры металлосодержащих комплексов (металлы изображены цветными кружками). *А*. Ca^{2+}-связывающий участок термолизина (Б. Мэтьюз). *Б*. Zn^{2+}-связывающий участок активного центра карбоксипептидазы А. Молекулы воды, расположенные вертикально, не показаны. [D.M.Blou, T.A.Steitz, Ann.Rev.Biochem., **39**, 78 (1970).]*В*. Высокопотенциальный железосерный белок *Chromatium*, содержащий четыре иона железа, четыре лабильных атома серы и четыре остатка цистеина. Сходную группу атомов имеют бактериальные ферредоксины (ср. рис. 2.28). [L. H. Jensen, Ann. Rev. Biochem. **43**, 461 (1974).] *Г*. Миоглобин. Железопорфириновый комплекс связан с белком через гистидин. На рисунке показано, что шестая координационная связь лиганда не занята, это — место связывания кислорода. Фактически идентичные структуры существуют и в гемоглобине. [R.E.Dickerson, I.Geis, The Structure and Action of Proteins, New York, Harper and Row (1969).]


Ala58
Asp57
H2O
H2O
Asp59
Ca
H2O
H2O
Asn60
Gln61
Phe62
A
170
His156
Zn
N
Glu72
O
N
His69
Б
His93
Cys43
Cys63
Cys77
Cys46
Fe
S
B
N
Fe
Г
His64

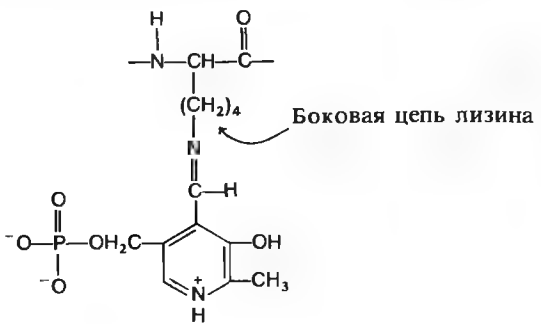


Биотин



Липоевая кислота

Протонированное шиффово основание ретиналя



Шиффово основание пиридоксальфосфата

РИС. 2.8. Структура некоторых простетических групп. Они дают представление о причинах появления у белков химических и фотохимических свойств, отсутствующих у обычных боковых групп аминокислот.

нений. Кроме того, во многих металлопротеидах возможна замена одного катиона на другой. Особенно информативно замещение атомов железа или цинка кобальтом, так как изучение различных физических и химических характеристик Co^{2+} позволяет определить, участвует ли катион в функционировании белка или его влияние оказывается более сложным. Одна из причин столь пристального внимания к металлопротеидам состоит в том, что атом металла можно использовать в качестве спектроскопического зонда, непосредственно входящего в функциональный центр фермента. Это значительно облегчает многие физические исследования, так как позволяет устранить нежелательные искажения структуры, вызываемые введением чуждых белку зондов.

Многие белки используют в качестве кофактора не металлы, а небольшие молекулы органических соединений. Большинство этих молекул связывается, по-видимому, с лизиновым остатком. Некоторые из них изображены на рис. 2.8. Все органические кофакторы сообщают белку химические свойства, которыми не обладают составляющие его обычные аминокислоты. Например, длинная гибкая цепь в биотине и в липоевой кислоте позволяет этим коферментам перемещать связывающийся с ними субстрат с одного места связывания в ферменте на другое. Детальное расположение этих органических молекул в белках неизвестно; неясно также, что происходит с локальной структурой в их присутствии. Тем не менее пиродоксали и ретиналь служат очень полезными спектроскопическими зондами, позволяющими получить информацию об их окружении в белках.

2.3. Первичная структура

Изучение ковалентной первичной структуры белка следует начать прежде всего с определения числа и типов полипептидных цепей, из которых он состоит, а также с установления размеров каждой из них. Во многих случаях отдельная изолированная белковая субъединица состоит из одной цепи, однако в природе распространены и многоцепочечные белки. Их возникновение может быть связано с расщеплением одной исходной цепи. В гл. 11 и 12 описаны методы, с помощью которых обычно можно разделить цепи в соответствии с их зарядом или молекулярной массой. Теми же способами можно приближенно определить их размеры. Иногда более точные размеры получают путем анализа химическим методом отношения числа концевых остатков к их общему числу или из данных об аминокислотном составе с учетом того, что количества наименее распространенных аминокислот в цепи должны выражаться целыми числами. Среди всех известных белков самые короткие цепи встречаются в некоторых полипептидных гормонах, например в секретине или глюкагоне, и в ряде «карликовых» белков, таких, как ферредоксин, содержащих от 25 до 100 аминокислот. Более типичны пептидные цепи, включающие от 100 до 500 аминокислот. Наибольшая из известных цепей состоит более чем из 3000 остатков, однако такие длинные цепи встречаются редко, если не считать фибриллярных белков.

Методы определения аминокислотной последовательности полипептидных цепей хорошо отработаны. На многих стадиях исследования применяются автоматизированные процедуры анализа, особенно при проведении поэтапного отщепления концевых аминокислотных остатков, начиная с N-конца. В благоприятных случаях можно определить последовательность 60 и большего числа остатков почти без вмешательства исследователя. Таким образом, для многих небольших белков определение аминокислотной последовательности является теперь едва ли не рутинной операцией. Для крупных белков или белков, обладающих плохой растворимостью, эта задача более сложна. Однако, как правило, если необходимо знать последовательность белка и есть готовность выполнить требующуюся работу, окончательный успех почти гарантирован. Относительная простота определения последовательности остатков сразу же вызывает два вопроса: какую информацию можно непосредственно извлечь из последовательности и как эти данные могут пригодиться при проведении других экспериментальных исследований?

ДИСУЛЬФИДНЫЕ И ДРУГИЕ ПОПЕРЕЧНЫЕ СВЯЗИ

Первый этап полного или даже частичного установления последовательности остатков обычно состоит в определении наличия каких-либо связей между различными частями цепи. Наиболее распространены дисульфидные связи, формирующиеся при окислении пар цистеинов. Образовавшийся дисульфид часто рассматривается как отдельный элемент и называется цистином. Как показано на рис. 2.9, система связей C—S—S—C в цистине нелинейна и обычно имеет неплоскую конформацию. Поэтому наличие цистина налагает сильные ограничения на белковую структуру. Два участка пептидной цепи должны быть сближены в пространстве и к тому же должны располагаться под определенными углами. Это означает, что при прочих равных условиях труднее разрушить структуру белков, содержащих дисульфидные связи. В самом деле, некоторые небольшие белки, в которых от 5 до 7 дисульфидов приходятся менее чем на 100 аминокислотных остатков, известны как наиболее стабильные. Они могут выдерживать экстремальные условия, возникающие при кипячении или воздействии сильных кислот, в которых очень многие белки довольно быстро и необратимо денатурируют.

Вообще говоря, большинство белков можно разделить на три класса в соответствии с их SH-группами. Небольшая часть белков не содержит ни цистеина, ни цистина; многие другие содержат один или несколько цистеинов, но не имеют цистинов. Существуют также белки с одним или чаще несколькими цистинами. И очень редко встречаются белки, содержащие и цистеины, и цистины. Для этого имеются достаточные основания, так как структура такой молекулы была бы нестабильной вследствие реакций дисульфидного обмена

$$—S—S—Cys_2 + Cys_3—SH \rightleftarrows Cys_1—SH + Cys_2—S—S—Cys_3$$

которые могут быть либо внутримолекулярными, либо межмолекулярными. Даже если свободные SH-группы и дисульфидные связи стерически не могут быть сближены, имеется лишь узкая область внешних условий, при которых они могут существовать в белковой молекуле одновременно. Единственным исключением являются сывороточные альбумины, которые обычно имеют 17 дисульфидных связей и одну свободную SH-группу. Однако эти белки весьма необычны также и в других отношениях.

Знание системы дисульфидных связей в белке иногда бывает полезным при определении его трехмерной структуры. Если белок состоит из нескольких цепей, то могут существовать как внутри-, так и межцепочечные дисульфидные связи и ковалентная структура обычно имеет довольно сложную топологию. В качестве примера на рис. 2.10 приведено схематическое изображение структуры иммуноглобулина G (IgG). Можно видеть 4 пептидные цепи и в общей сложности 15 дисульфидных связей — 4 межцепочечные, а остальные внутрицепочечные. Топологическая модель, показанная на этом рисунке, позволяет прийти к интересному предположению, что данный белок состоит из областей, каждая из которых независимо скрепляется своими дисульфидами. Чем больше появляется точных структурных данных об иммуноглобине G, тем яснее становится, что белок представляет собой совокупность отдельных доменов.

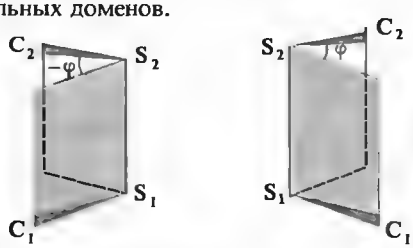


РИС. 2.9. Структура цистиновых дисульфидных связей. Отметим, что метиленовые группы расположены в разных плоскостях. ϕ — двугранный угол между плоскостями. Показаны два оптических изомера с одним и тем же значением угла ϕ. (По рисунку, предоставленному С. Бейхоком.)

Кроме дисульфидных связей в белках было обнаружено несколько других типов поперечных связей. Они, однако, встречаются значительно реже. Некоторые фибриллярные белки, особенно часто присутствующие в соединительных тканях, содержат ряд межцепочечных сшивок. Последние служат для увеличения жесткости, а также, возможно, обеспечивают обратимость деформации структуры белка. Большинство этих сшивок возникает при превращении ε-аминогруппы лизина в альдегидную под действием лизилоксидазы. В ходе этой реакции образуются уникальные функциональные группы, которые могут либо димеризоваться посредством альдольной конденсации с образованием поперечных связей между остатками, либо реагировать с немодифицированным лизином. Каждая такая ре-

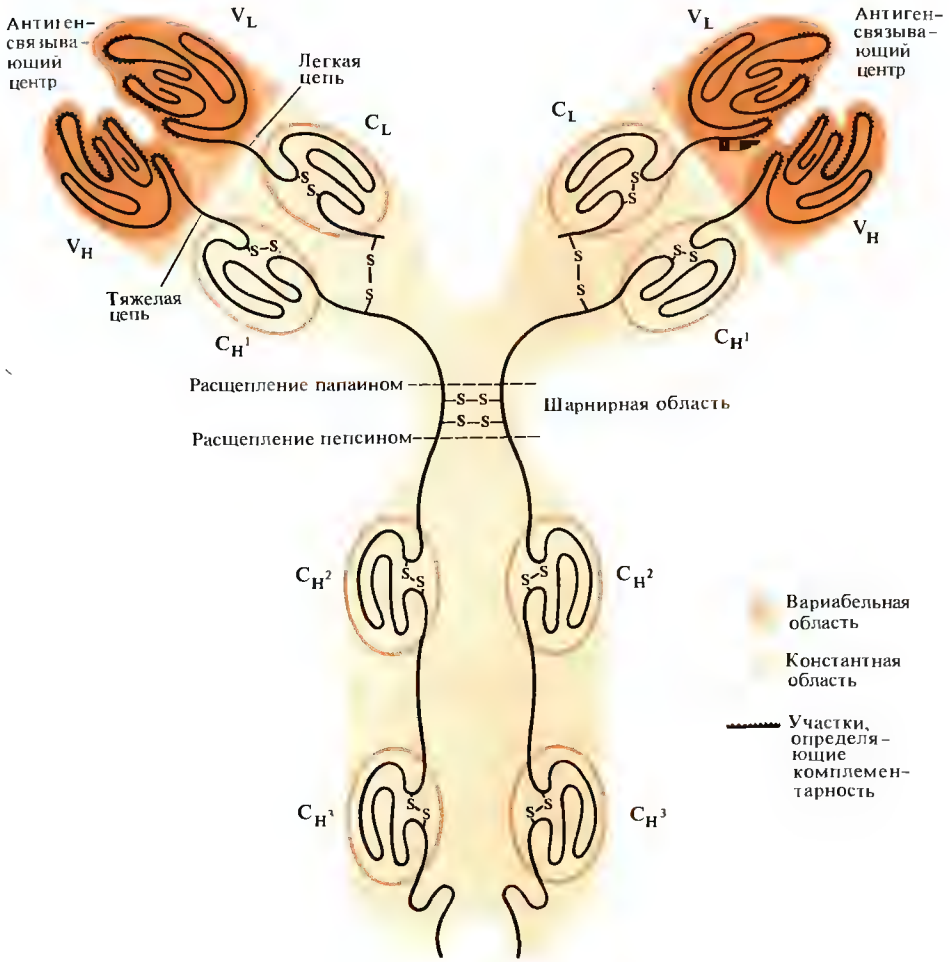


РИС. 2.10. Модель дисульфидных связей в иммуноглобулине G. Индексы *L* и *H* отиосятся к легким и тяжелым цепям соответственно, буквы *C* и *V* обозначают области, аминокислотиая последовательность которых сохраняется или, наоборот, значительно меняется у различных видов IgG. Каждая выделенная область последовательностн сворачивается в домен с независимой третичной структурой, и, таким образом, результирующая структура напоминает белок, состоящий из 12 субъединиц, хотя он представляет собой одну ковалентную единицу (рис. Ирвинга Гейса).

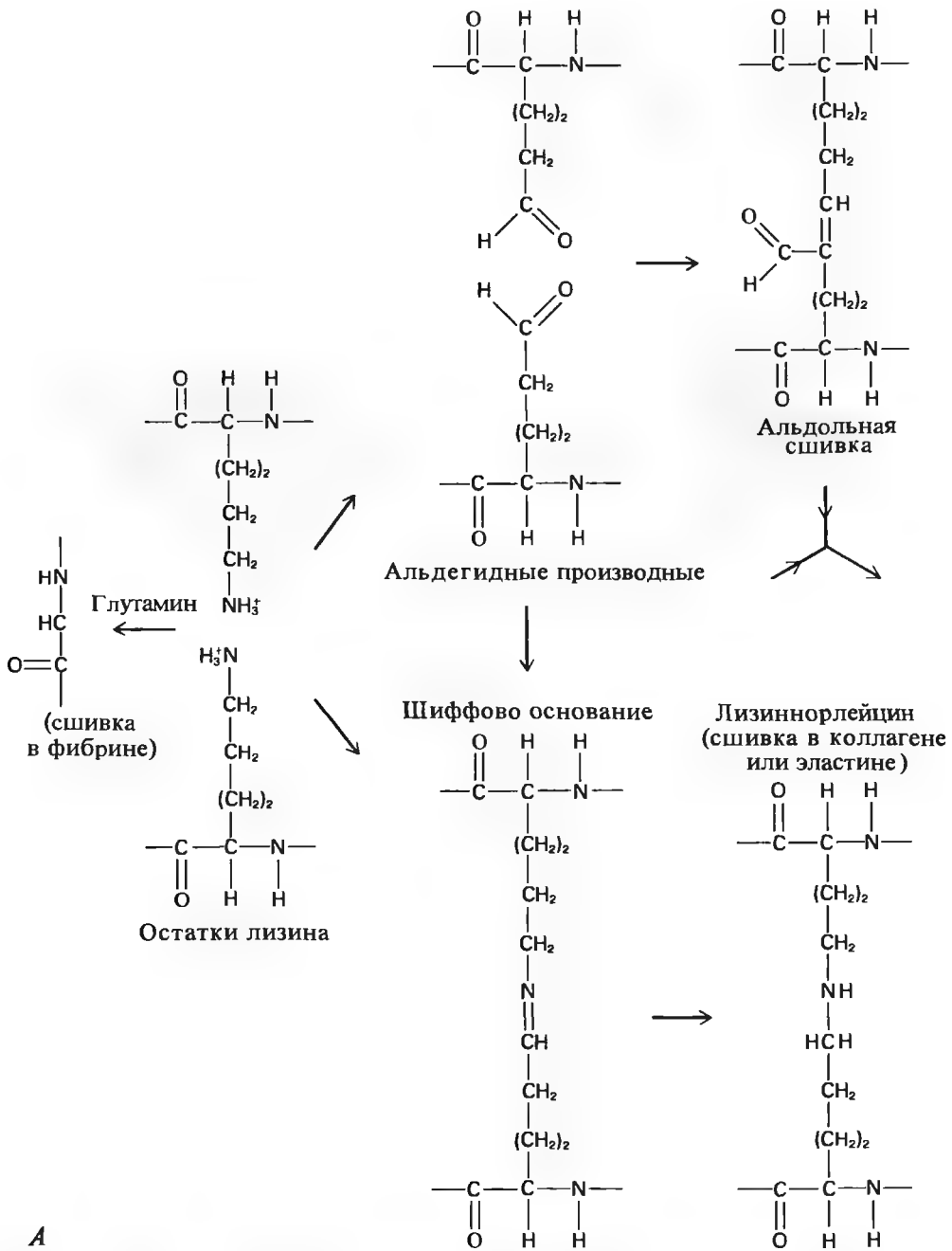
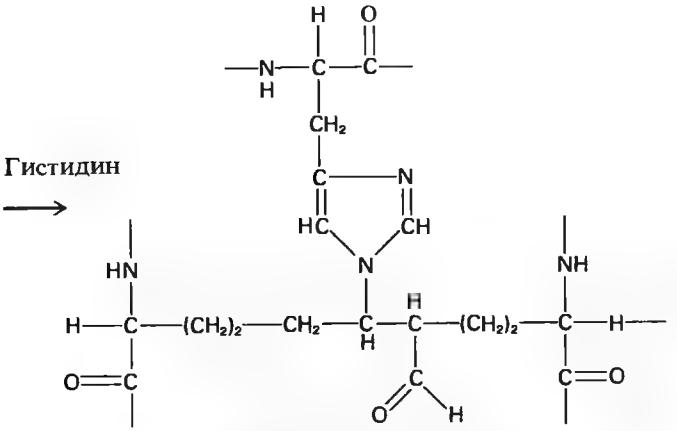
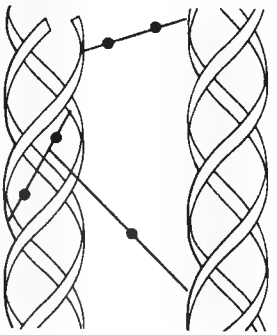


РИС. 2.11. Некоторые пути образоваиия лизиновых сшивок в коллагене, эластине и фибрине. *А*. Лизины могут соединяться друг с другом, с гистидином или с глутамином, образуя соответствеино димер, тример и тетрамер. *Б*. Схематическая иллюстрация того, как внутримолекулярные и межмолекулярные связи могут скреплять между собой молекулы в волокнах коллагена.

Гистидин

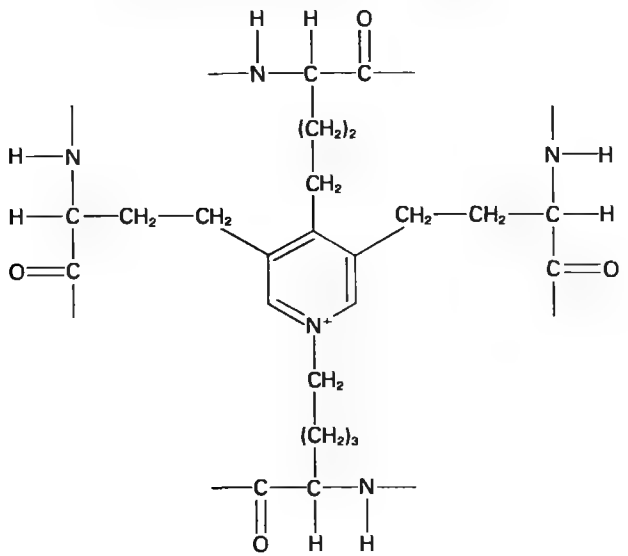


Гистидин-альдольная сшивка



Б

Десмозин (сшивка в эластине)



акция может в свою очередь порождать цепочку новых реакций, которые часто приводят к объединению трех или четырех аминокислотных боковых цепей. Некоторые из известных поперечных лизиновых связей в коллагене и эластине показаны на рис. 2.11. На том же рисунке показаны сшивки между глутамином и лизином, которые встречаются в фибрине, белке, участвующем в свертывании крови. Эти поперечные связи, по-видимому, стабилизируют сгустки крови.

ПЕРВИЧНАЯ СТРУКТУРА И АНАЛИЗ ВТОРИЧНОЙ И ТРЕТИЧНОЙ СТРУКТУРЫ

Знание первичной структуры необходимо для многих биохимических и биофизических исследований. Ярким примером этого является определение трехмерной структуры белка при помощи рентгеноструктурного анализа. Этот метод редко обладает достаточно высоким разрешением, чтобы можно было увидеть отдельные атомы. Часто рентгеноструктурные данные оказываются недостаточно ясными даже для того, чтобы можно было однозначно проследить ход полипептидной цепи по всей структуре белка. Однако при обычном разрешении, как правило, просматривается большая часть полипептидного скелета, а некоторые большие боковые группы видны довольно отчетливо. В таких случаях одновременное использование данных об аминокислотной последовательности и кристаллографических карт электронной плотности может позволить найти единственно возможную структуру, соответствующую тем и другим данным. Более подробно этот метод рассматривается в гл. 13, но необходимо помнить, что обычно метод рентгеноструктурного анализа не обладает достаточным разрешением, чтобы была видна вся последовательность аминокислот. Следует также учитывать, что данные об этой последовательности сами часто содержат ошибки. Совместное использование обоих методов значительно расширяет их возможности.

Не менее важна роль данных о первичной структуре при анализе результатов, полученных любыми методами, использующими химическую модификацию или введение в молекулу метки. Если известно только то, что сшивающий реагент провзаимодействовал с остатками лизина или что спектроскопическая метка присоединилась именно к лизину, то это еще мало о чем говорит. Можно ли считать образовавшийся продукт единственным? Дает ли это какую-либо информацию об укладке пептидной цепи с образованием трехмерной структуры? Зная последовательность и идентифицировав с помощью пептидных карт меченые пептиды, можно определить специфические места химической модификации.

Например, на рис. 2.12 показан полипептидный скелет бычьей рибонуклеазы А. Химическое исследование белка показывает, что 1,5-дифтор-2,4-динитробензол сшивает остатки лизина 7 и 41. Неожиданная чувствительность белка к модификации иодуксусной кислотой приводит к любопытному заключению: гистидины 12 и 119 модифицируются по принципу «или-или»; это можно объяснить тем, что они располагаются в активном центре фермента недалеко друг от друга. Сопоставляя сведения о последовательности и обе группы химических данных, легко видеть, что плоское расположение пептидной цепи исключается. Молекула должна быть уложена в пространстве более сложным образом, и результаты рентгеноструктурного анализа действительно подтверждают это.

ПЕРВИЧНАЯ СТРУКТУРА И ПРЕДСКАЗАНИЕ ВТОРИЧНОЙ И ТРЕТИЧНОЙ СТРУКТУРЫ

Иногда анализ последовательности аминокислот может привести к более определенным выводам о структуре белка. В настоящее время существуют методы, позволяющие с некоторой (хотя и не очень большой) степенью достоверности предсказывать расположение разных типов вторичной структуры вдоль цепи. Они подробно обсуждаются в гл. 5.

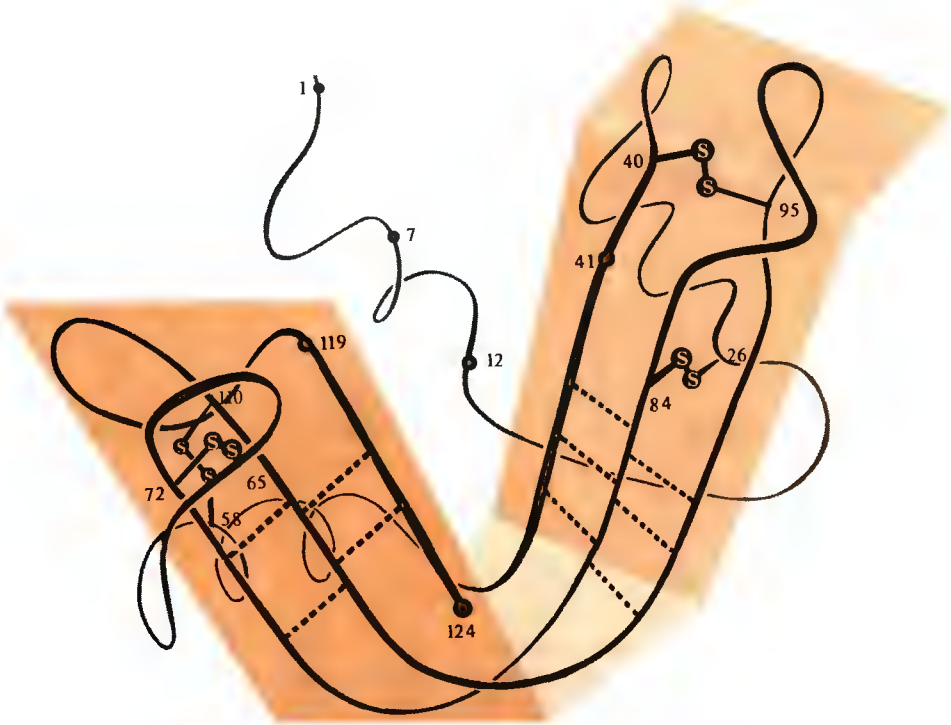


РИС. 2.12. Пептидный остов рибонуклеазы из поджелудочной железы быка. Показаны две пары остатков, которые располагаются в третичной структуре вблизи друг друга или потому, что они скреплены сшивкой (Lys^{7} и Lys^{41}), или потому, что они непосредственно участвуют в функционировании активного центра (His^{12} и His^{119}). (Рисунок Ирвинга Гейса.)

Современное состояние этого вопроса иллюстрирует рис. 2.13, где сравниваются результаты применения нескольких различных подходов к определению структуры аденилаткиназы. Была предпринята попытка предсказать по аминокислотной последовательности расположение вдоль цепи участков вторичной структуры трех типов: α-спиралей, β-слоев и β-изгибов. Результаты сравнивались с действительной трехмерной структурой, полученной методом рентгеноструктурного анализа. Как видно из рис. 2.13, расположение α-спиралей и β-изгибов довольно хорошо предсказывается несколькими различными методами, тогда как точность локализации β-слоев все еще оставляет желать лучшего.

Наиболее заманчиво использовать информацию об аминокислотной последовательности для предсказания третичной структуры белковой молекулы, а отсюда, возможно, и ее функции. Некоторые принципы и положения, применяемые для решения этой задачи, изложены в гл. 5. Здесь мы приведем один результат, который иллюстрирует современное состояние этой области исследования. На рис. 2.14 проведено сравнение трехмерной структуры ингибитора трипсина из поджелудочной железы быка, определенной методом рентгеноструктурного анализа, с модельной структурой, построенной на основании данных об аминокислотной последовательности, в которых использовалась информация о термодинамике взаимодействий между аминокислотными остатками. Исходной считалась вытянутая конформация. Чтобы проследить процесс укладки молекулы, рассчитывали силы, действующие между различными остатками, полученные из данных об энергии их взаимодействия. В результате достигли неплохого качественного согласия между

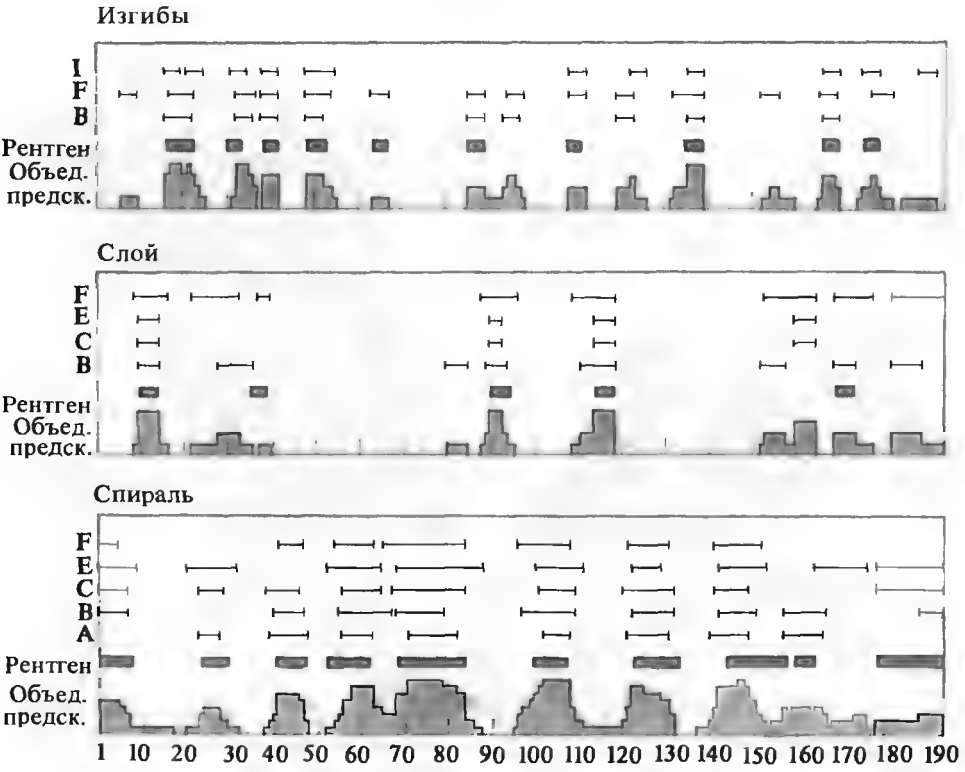


РИС. 2.13. Распределение типов вторичной структуры в аденилаткиназе. Указаны области, которые, согласно предсказаниям различных теоретических методов, считаются β-слоями, β-изгибами или α-спиралями. На рисунке представлены также области вторичной структуры, определенные методом рентгеноструктурного анализа, и результаты, полученные при объединении всех теоретических подходов. Предсказания были сделаны до того, как стала известна структура белка. [G.E.Shultz et al., Nature, **250**, 140 (1974).]

экспериментом и теорией. Расчеты действительно предсказали образование глобулярной структуры. Более того, полученное в результате вычислений расположение некоторых участков пептидного скелета оказалось весьма сходным с действительным. Это предсказание получилось не столь успешным, чтобы можно было вообще отказаться от трудоемких методов рентгеноструктурного анализа белков. Все же результаты расчетов достаточно близки к реальности, чтобы с оптимизмом относиться к перспективе развития подобных теоретических подходов.

Некоторые последовательности аминокислот содержат повторы. Как показано в примерах на рис. 2.15, эти области могут быть различными по протяженности; это сразу наводит на мысль о возможности существования в трехмерной структуре элементов симметрии или псевдосимметрии. Длинные повторы могут привести к образованию в молекуле либо симметричных областей, либо множественных доменов, которые могут быть похожими на субъединицы субъединичных белков. Короткие повторы могут указывать на наличие спиральных сверхспиральных структур — типа структур, обнаруженных в коллагене и других фибриллярных белках.

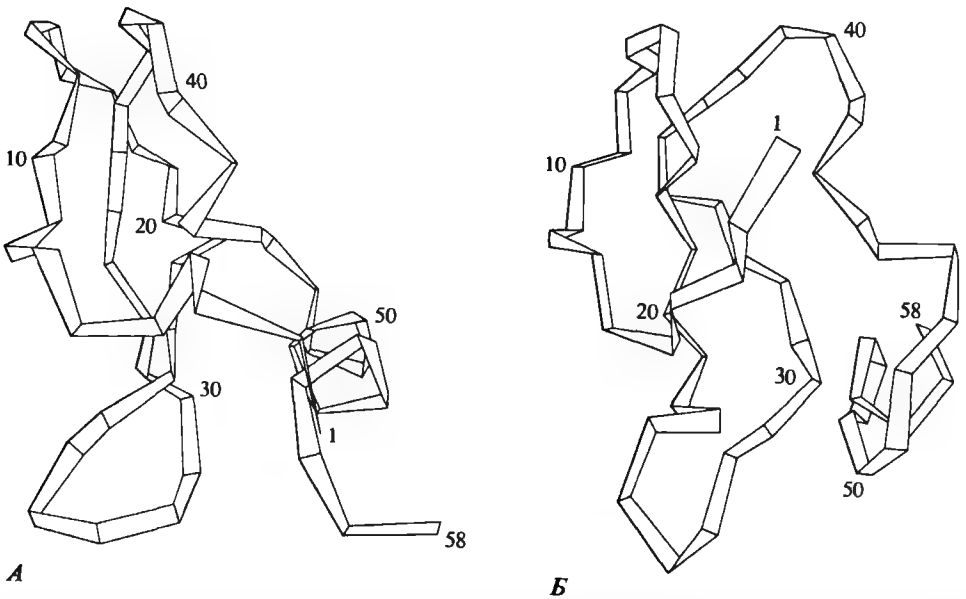
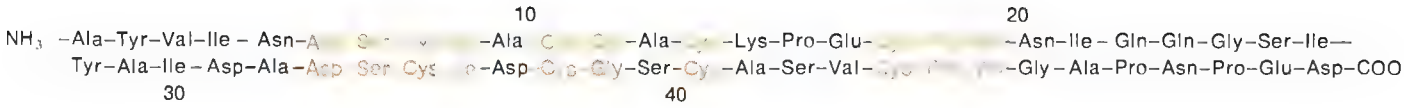


Рис. 2.14. Схематическое изображение пептидного остова ингибитора трипсина из поджелудочной железы. *А*. Изображение структуры, полученной методом рентгеноструктурного анализа. *Б*. Структура, найденная методом минимизации конформационной энергии молекулы, причем в качестве начальной принималась развернутая цепь (за исключением спирали, расположениой иа С-конце). [M.Levitt, A.Warshel, Nature, 293, 693 (1975).]

АМИНОКИСЛОТНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ БЕЛКА И АНАЛИЗ ЕГО ФУНКЦИЙ

Данные об аминокислотной последовательности могут оказаться чрезвычайно информативными в тех случаях, когда что-то уже известно о функции белка, особенно если он взаимодействует с другими крупными биополимерами. Два примера белков такого рода (гистон и гликофорин) приведены на рис. 2.16. Гистоны прочно связываются с ДНК. Мы уже указывали, что большое число положительно заряженных аминокислот в гистонах, по-видимому, необходимо для стабилизации взаимодействия с таким полианионом. Как видно из рис. 2.16, *А*, у N-конца последовательности находится необычный кластер положительно заряженных аминокислот. Напрашивается предположение, что этот участок имеет вытянутую конформацию и может спиралеобразно закручиваться вдоль довольно протяженного участка ДНК. Остальная часть белковой молекулы имеет обычный аминокислотный состав и, по-видимому, свертывается с образованием типичной глобулярной структуры.

Гликофорин — мембранный белок, выделяемый из эритроцитов. Химические данные наводят на мысль, что он является трансмембранным белком, т.е. одна его часть находится вне клетки, другая — внутри, а третья как бы «прошивает» мембрану. Известно, что остатки сахаров мембранных гликопротеидов почти всегда находятся на наружной



А

Б

					5				10					15
1	Gly	Tyr	Asp	Glu	Lys	Ser	Ala	Gly	Val	Ser	Val	Pro	Gly	Pro	Met
16	Gly	Pro	Ser	Gly	Pro	Arg	Gly	Leu	Pro	Gly	Pro	Pro	Gly	Ala	Pr
31	Gly	Pro	Gln	Gly	Phe	Gln	Gly	Pro	Pro	Gly	Glu	Pro	Gly	Glu	Pr
46	Gly	Ala	Ser	Gly	Pro	Met	Gly	Pro	Arg	Gly	Pro	Pro	Gly	Pro	Pro
61	Gly	Lys	Asn	Gly	Asp	Asp	Gly	Glu	Ala	Gly	Lys	Pro	Gly	Arg	Pro
76	Gly	Gln	Arg	Gly	Pro	Pro	Gly	Pro	Gln	Gly	Ala	Arg	Gly	Leu	Pr
91	Gly	Thr	Ala	Gly	Leu	Pro	Gly	Met	Lys	Gly	His	Arg	Gly	Phe	Ser
106	Gly	Leu	Asp	Gly	Ala	Lys	Gly	Asn	Thr	Gly	Pro	Ala	Gly	Pro	Lys

РИС. 2.15. Аминокислотная последовательность двух белков, включающая повторы. *А*. Ферредоксин *P. aerogenes*; мы видим значительное сходство первых и последних 27 аминокислот. Восемь цистеиновых групп образуют координационные связи с атомами железа (см. рис. 2.7 и рис. 2.8). *Б*. Фрагмент коллагена. За исключением трех остатков вблизи N-конца каждый третий остаток является глицином. Большое число пролиновых остатков не допускает образования α- и β-структур и приводит к формированию спирали полипролинового типа. Периодическое расположение остатков глицина способствует образованию трех спиралей, которые, в свою очередь переплетаясь, образуют трехцепочечную коллагеновую спираль.

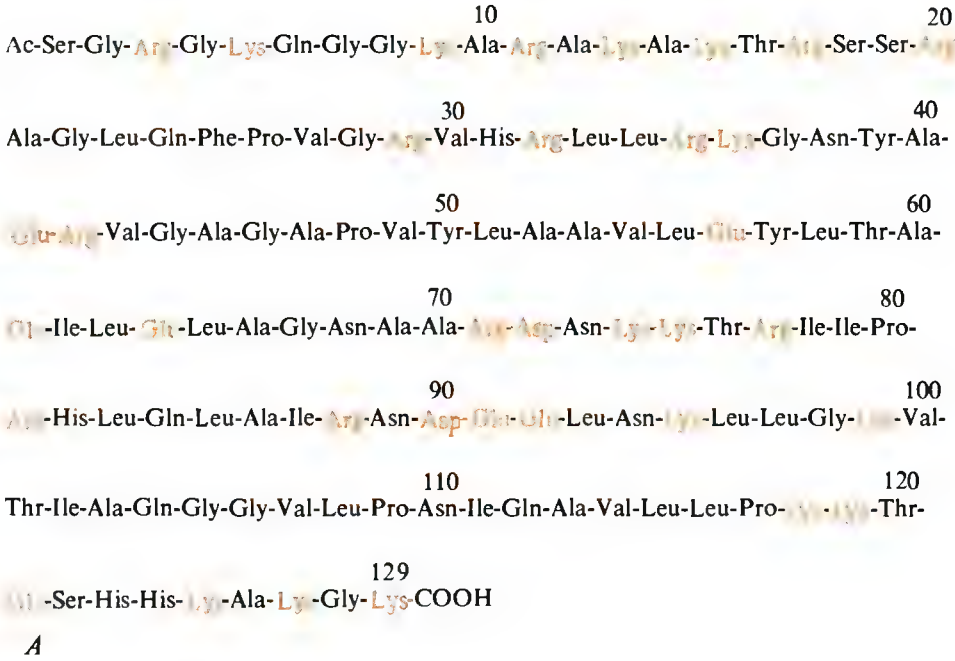
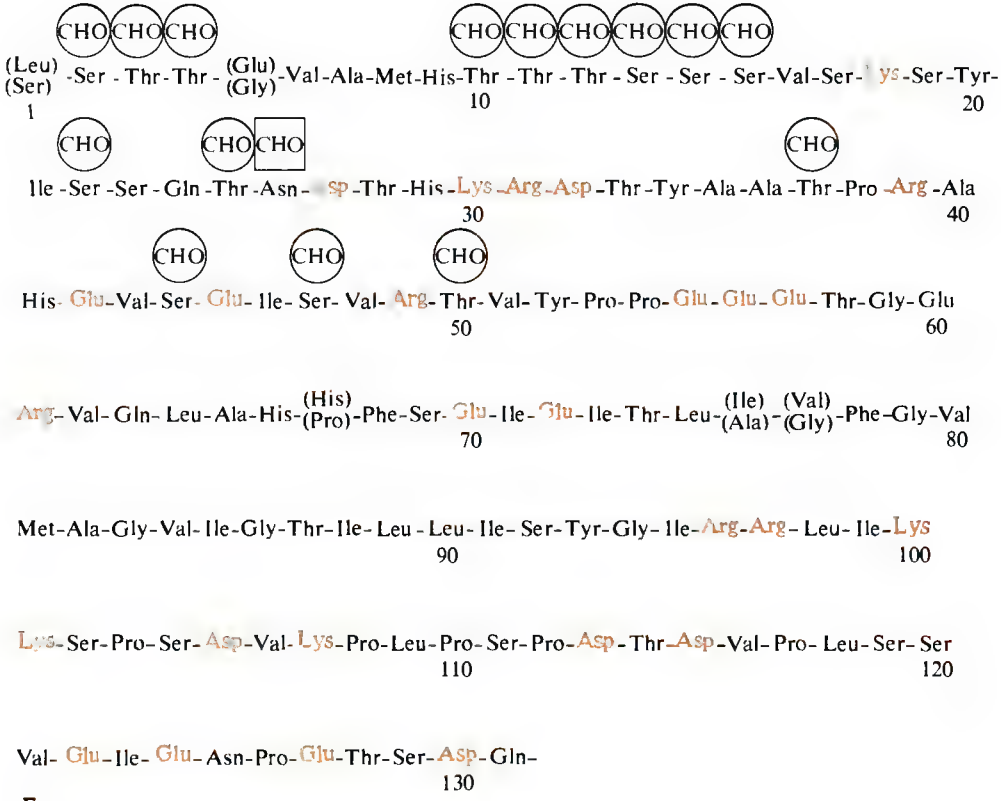


РИС. 2.16. Аминокислотные последовательности двух белков, обнаруживающие некоторые особенности структуры и функции. (Все заряженные остатки показаны цветными буквами.) *А*. Гистон 2а теленка. Первые 36 остатков содержат 12 положительных зарядов и ни одного отрицательного. Эта область, по-видимому, взаимодействует с отрицательно заряженными фосфатами двойной спирали ДНК.



Б

РИС. 2.17. *Б*. Гликофорин эритроцитов человека. CHO обозначает остатки, к которым олигосахариды присоединяются или О-гликозидной связью (кружок), или N-гликозидной связью (квадратик). Гликофорин является трансмембранным белком. Сахара должны располагаться на внешней стороне клетки. Длинный участок, состоящий из неполярных остатков от Ile^{73} до Ile^{95}, по-видимому, проникает через мембрану, и поэтому есть основания предполагать, что С-конец находится внутри клетки. [M.Tomita, V.T.Marchesi, Proc.Natl.Acad.Sci. USA, **72**, 2914 (1976).]

поверхности клетки. Поскольку эти остатки сконцентрированы на N-концевой половине последовательности гликофорина, именно данная часть белковой молекулы должна быть внешней. Тот участок гликофорина, который находится внутри мембраны, должен быть неполярным (гидрофобным), чтобы мог образоваться тесный контакт с липидами. При исследовании первичной структуры гликофорина выявляется в высшей степени необычный участок из 20 аминокислотных остатков, имеющих достаточно высокое сродство к липидам. Если бы этот участок находился в развернутой конформации, цепочка оказалась бы достаточно длинной, чтобы протянуться через весь липидный бислой. В таком случае короткая С-концевая полярная область должна находиться внутри клетки. Заметим, что переход к полярным остаткам с обеих сторон неполярной области довольно резок. Можно предположить, что молекула жестко фиксирована в липидном бислое и что любые ее перемещения вне или внутрь клетки связаны с большими энергетическими затратами.

СРАВНЕНИЕ АМИНОКИСЛОТНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ РОДСТВЕННЫХ БЕЛКОВ

В гл. 1 говорилось, что данные об аминокислотной последовательности становятся более убедительными и информативными, когда мы имеем возможность сравнивать последовательности структурно и функционально родственных молекул. Здесь приводится несколько таких примеров для белков. Наиболее полная информация об аминокислотных последовательностях накоплена для гемоглобинов и родственных им гемсодержащих белков — миоглобинов. Сравнение последовательностей в пределах одного вида с очевидностью указывает на то, что между α- и β-субъединицами гемоглобинов и миоглобином существует четкое соответствие. По-видимому, у какого-то предкового организма был только один глобиновый ген, который после многократных дупликаций оказался представленным в виде нескольких копий. В результате дивергенции этих новых генов в ходе эволюции и появилось все множество глобинов. Анализируя последовательности либо «на глаз», либо более тщательно с помощью компьютера, можно построить генеалогическое древо или, говоря эволюционным языком, филогенетическое древо. Миоглобин в большей степени отличается от α- и β-цепей гемоглобина, чем эти цепи друг от друга. Следовательно, ген миоглобина, вероятно, должен был обособиться от гена гемоглобина до дивергенции α- и β-цепей (рис. 2.17).

Вместо β-цепей в гемоглобине человеческого плода содержатся две γ-цепи. Сравнение последовательностей показывает, что γ-цепь значительно больше похожа на β-цепь, чем на α-цепь гемоглобина взрослого человека. В гемоглобине взрослого человека присутствует в небольшом количестве еще один представитель семейства гемоглобинов, имеющий две δ-цепи вместо β-цепей. Они еще больше похожи на β-цепь. Таким образом, у человека есть по меньшей мере три гена на каждую гемоглобиновую цепь β-типа в гемоглобине. Поскольку все эти белки являются функциональными, имеет смысл сравнить между собой их последовательности. Вероятно, изменяющиеся от последовательности к последовательности остатки не являются для белка функционально важными. Сравнительный анализ можно расширить, сопоставляя цепи гемоглобина человека и других позвоночных. Можно предположить, что при дивергенции организмов в ходе эволюции различие последовательностей аминокислот в гемоглобине также будет увеличиваться. Действительно, гемоглобин человека по своей аминокислотной последовательности очень близок к гемоглобину гориллы, но значительно отличается от гемоглобина кенгуру и совершенно не похож на гемоглобин лягушки или карпа.

Существуют два способа оценки полной картины аминокислотных последовательностей, учитывающей их структуру и функции. Человек, склонный к обобщениям, будет утверждать, что все эти белки, будучи гемоглобинами, выполняют одинаковые функции. Сравнивая аминокислотные последовательности в гемоглобинах многих биологических видов, можно заметить, что только часть остатков тождественна для всех видов

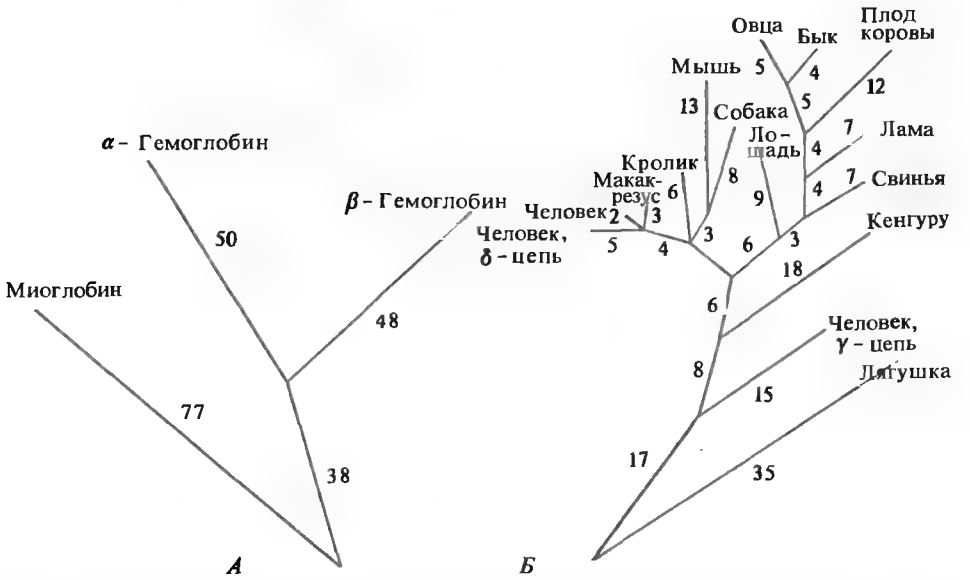


РИС. 2.17. Филогенетическое древо, построенное путем сравнения последовательностей аминокислот в гемоглобинах и миоглобине. Длина каждой ветви пропорциональна числу точечных мутаций, необходимых для перехода одной последовательности в другую. Точки ветвления обозначают положения возможных общих предков. Числа на каждой ветви древа указывают число мутаций, приходящееся на каждые 100 остатков. *А*. Полное семейство глобинов кенгуру. *Б*. Подробное представление β-цепей гемоглобиов и белков родственных ему типов у различных видов. [M.O.Dayhoff et al. Atlas of protein sequence and structure, vol. 5 (Silver spring, Md.: National Biomedical Research Foundation, 1972).]

(рис. 2.18). Константные области, вероятно, абсолютно необходимы для обеспечения функции гемоглобина. На самом деле это не так просто, как кажется, и в функционировании этого белка принимают участие и другие структуры. По-видимому, во многих случаях пары (или бóльшие группы) аминокислот меняются таким образом, что при этом третичная структура белка остается прежней, несмотря на изменение его первичной структуры.

Распространяя это обобщение далее, можно заключить, что α- и β-цепи имеют почти одинаковую структуру. При сравнении аминокислотных последовательностей обеих цепей число идентичных функционально необходимых остатков еще в большей степени уменьшается. Оно может стать еще меньше, если учесть также последовательности в миоглобинах. Здесь надо, однако, иметь в виду, что гемоглобин функционирует как агрегат субъединиц, тогда как миоглобин агрегатов не образует. Поэтому уменьшение числа неизменных остатков при совместном анализе последовательности α- и β-цепей и миоглобина может пролить свет на то, какие остатки важны для контактов субъединиц в гемоглобинах.

Ученые, интересующиеся деталями механизма функционирования гемоглобина, рассматривают филогенетическое древо с другой точки зрения. Хотя все эти белки — гемоглобины, требования, предъявляемые к ним различными организмами, неодинаковы. Например, холоднокровные и теплокровные организмы существуют при различных темпе-

ратурах; организмы, обитающие на большой высоте, нуждаются в особых свойствах кислородсвязывающих белков; рыбам нужно наполнять свой плавательный пузырь и т.д. Аминокислотные последовательности как раз и могут стать золотой жилой при изучении процессов перестройки структуры белка, происходящих в ответ на индивидуальные потребности различных организмов.

Другое семейство близкородственных белков образует несколько сериновых эстераз. К ним относятся протеолитические ферменты химотрипсин, трипсин, эластаза и тромбин. На рис. 2.19 сравниваются аминокислотные последовательности этих четырех белков. Сравнение выявляет соответствие не только в аминокислотной последовательности, но и в расположении многих дисульфидных поперечных связей, а также в локализации очень реакционноспособного остатка серина, который, как известно, находится в активных центрах всех этих ферментов. Можно предположить, что такое сходство первичных структур должно приводить к сходству их третичной структуры. Именно это и представлено на рис. 2.20, где изображены три из четырех упомянутых выше белков. Следует, однако, обратить внимание на то, что, несмотря на сходство последовательностей, структуры и механизмов функционирования, позволяющее рассматривать эти четыре белка как родственные в эволюционном смысле, все же считать их тождественными никак нельзя. Различием аминокислотных последовательностей, и особенно пространственных структур, можно объяснить некоторые особенности субстратной специфичности этих белков и механизма их действия.

МУТАНТНЫЕ БЕЛКИ; ГЕМОГЛОБИН

Основываясь на результатах анализа аминокислотных последовательностей мутантных белков, можно попытаться идентифицировать остатки, играющие важнейшую роль в функционировании белков. Поскольку большинство мутаций приводит к определенным функциональным нарушениям, их изучение дает возможность понять специфические эффекты аминокислотных замен в белке. Однако выбор мутантных белков, с которыми можно работать, несколько ограничен. Если мутация доминантна и летальна, то (по определению летальной мутации) организм не способен давать потомство или даже жить достаточно долго, чтобы можно было выделить мутантный белок. Поэтому наиболее вредные мутации реально наблюдать нельзя. Когда дело касается бактерий, то имеются методы, позволяющие обойти эти трудности, например использование температурозависимых мутаций, которые летальны при одних температурах, но допускают развитие при других. К сожалению, проводить генетические эксперименты на высших организмах как с практической, так и с этической точки зрения значительно труднее.

В результате многочисленных исследований обнаружено несколько сотен мутантных гемоглобинов человека. На рис. 2.18 суммированы данные для β-субъединиц ряда этих мутантных форм. При тщательном изучении рисунка видно, что мутации локализованы в участках структуры, в которых меньше межвидовых различий. Это можно было предсказать заранее, если считать, что определенные остатки необходимы для нормального функционирования гемоглобина.

Подробный анализ фенотипа каждой мутантной формы очень полезен для понимания молекулярных основ многих болезней человека, однако он выходит за рамки этой книги. Особого упоминания заслуживает лишь гемоглобин S. Он был первым обнаруженным мутантным гемоглобином. В нем происходит замена всего одной аминокислоты в шестом положении β-цепи (глутаминовая кислота заменяется на валин). Этот мутантный белок интересен тем, что он в отличие от остальных широко распространен среди жителей Центральной Африки. У гомозигот по гену гемоглобина S развивается особая болезнь — серповидноклеточная анемия. В условиях пониженного парциального давления кислорода в крови, особенно в капиллярах, у носителя такого гена изменяется форма эритроцитов — они превращаются из двояковогнутых дисков в вытянутые нерегулярные структуры. Такие «серповидные» клетки, задерживаясь в капиллярах, препятствуют нор-

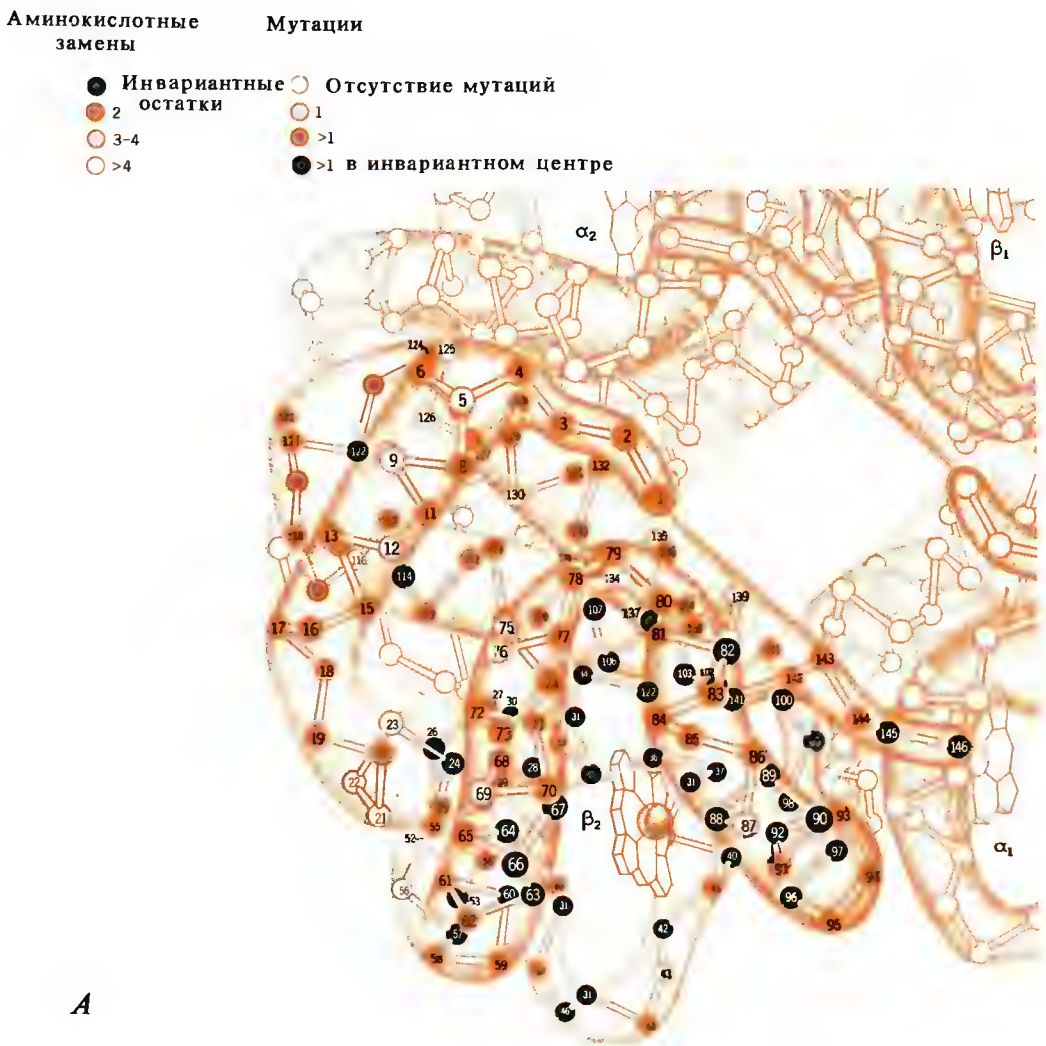
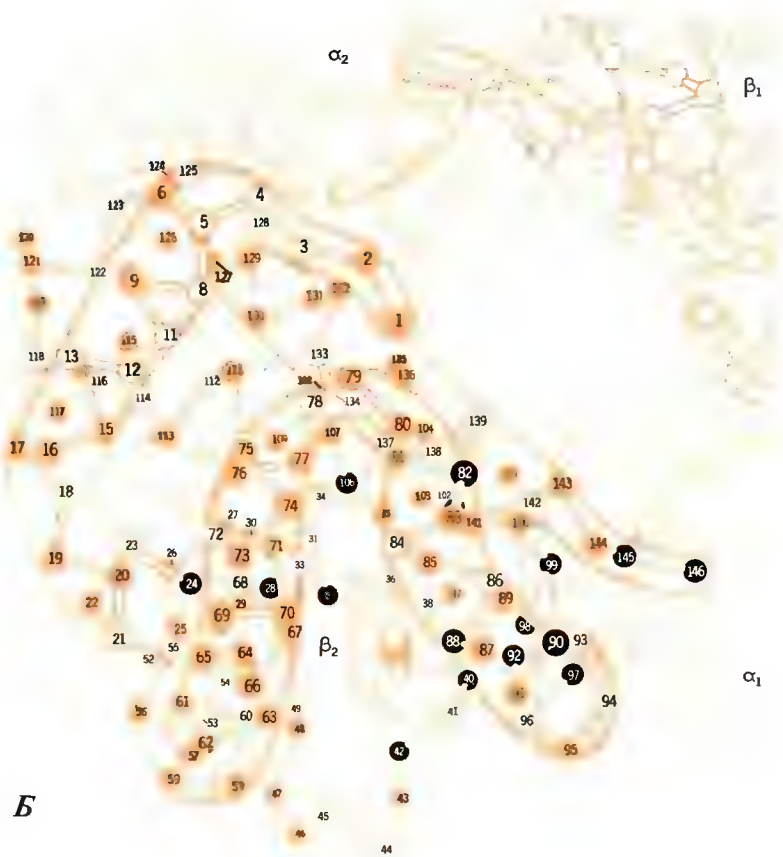


РИС. 2.18. Структура β-цепи гемоглобина человека, иллюстрирующая контакты с соседними субъединицами. *А*. Показаны остатки, которые могут варьировать в β-подобных цепях обычных гемоглобинов позвоночных. Заметим, что большинство инвариантных остатков располагается вблизи кармана, связывающего гем, или в области контакта α_1- и β_2-субъединиц, где в процессе оксигенации происходят изменения четвертичной структуры. Очевидно, эти области имеют важные структурные особенности, требующие уникальности остатков. С другой стороны, тесный контакт α_2- и β_2-субъединиц остается в процессе оксигенации практически неизменным. Здесь возможен выбор из



большого числа вариантов, сохраняющих в целости этот контакт. *Б*. Показана локализация некоторых мутаций в β_2-цепи гемоглобина человека. Они сосредоточены в области гемового кармана и на границе раздела α_1 — β_2-субъединиц. Это согласуется с представлением о том, что структура данных областей крайне важна для нормальной функции белка. Было замечено, что многие остатки в β-цепях гемоглобина не подвергаются никаким изменениям и мутациям. Отчасти это отражает статистическую природу мутагенеза, однако некоторые остатки, по-видимому, настолько важны, что любое их изменение приводит к летальному исходу. (Рисунки Ирвинга Гейса.)

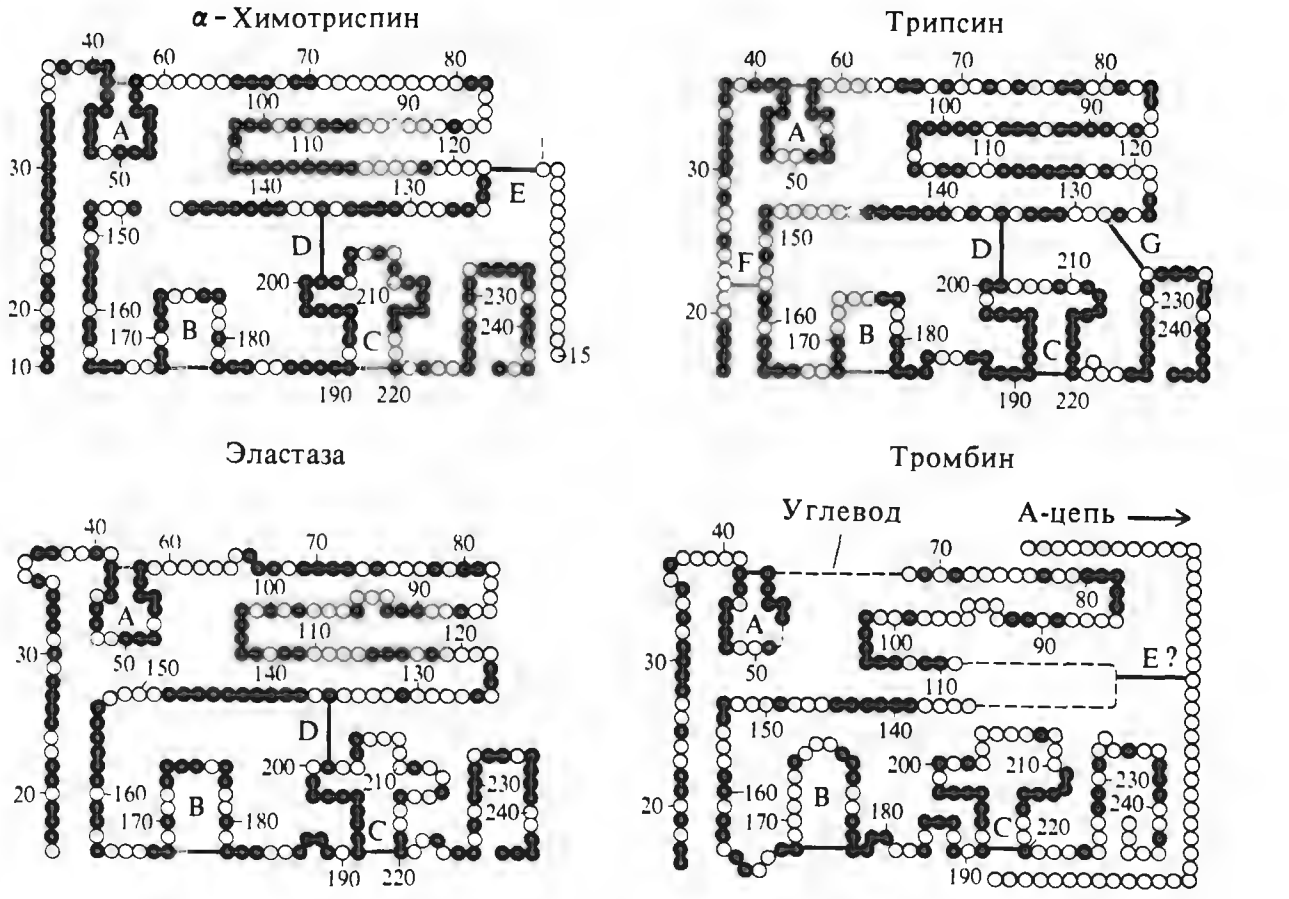


РИС. 2.19. Схематическое изображение аминокислотных последовательностей четырех родственных белков из семейства сериновых эстераз. Черные кружки означают остатки, тождественные по крайней мере в двух белках. Буквами от *A* до *F* обозначены соответствующие дисульфидные мостики (они показаны красными черточками). [B.S.Harley, Phil. Trans. Roy. Soc. London, Ser. B, **257**, 77 (1970).]

мальной циркуляции крови, что влечет за собой серьезные повреждения тканей. Кроме того, будучи более хрупкими, серповидные эритроциты разрушаются быстрее, и это приводит к развитию анемии.

Каким же образом простая замена одной аминокислоты может привести к таким большим изменениям в структуре целой клетки? Отметим, что в гемоглобине аминокислотный остаток $\beta 6$ находится вдали от группы гема (см. рис. 2.18). Он также не входит в область поверхности раздела $\alpha_2\beta_2$-субъединиц тетрамера. Поэтому связывание кислорода гемоглобином S не должно существенно отличаться от нормального. Однако эта мутация приводит к появлению неполярного остатка на внешней стороне молекулы, там, где в нормальной молекуле обычно располагается отрицательно заряженный остаток. Даже если больше ничего не меняется, можно думать, что такая замена повышает склонность молекул гемоглобина к агрегации. Трудно, правда, предсказать a priori, что гемоглобин в дезоксиформе должен образовывать строго упорядоченные трубчатые агрегаты, показанные на рис. 2.21. Эти трубки имеют тенденцию образовывать кластеры, ориентируясь параллельно друг другу. Чтобы разместились такие структуры, эритроциты должны изменить свою форму. Большая ось серповидной клетки параллельна большим осям содержащихся в ней трубчатых кластеров. Некоторые детали молекулярного механизма серповидноклеточной анемии приведены в Дополнении 2.2.

Дополнение 2.2

ИССЛЕДОВАНИЯ ВОЛОКОН, ФОРМИРУЕМЫХ ДЕЗОКСИГЕМОГЛОБИНОМ S

Информацию о структуре волокон гемоглобина S(HbS) дают экспериментальные данные трех типов. Электронная микроскопия и метод дифракции рентгеновских лучей на волокиах позволяют получить представление об упаковке тетрамеров гемоглобина, изображенной на рис. 2.21,*А*. Однако имеющегося разрешения недостаточно, чтобы различить ориентацию отдельной $\alpha_2\beta_2$-субъединицы в волокне. Это делает тщетным попытки выяснить, существует ли непосредственный контакт между мутантным $Val^{\beta 6}$ в одном тетрамере и участками поверхности соседних тетрамеров в трубчатом волокне. Если бы такие участки были известны, можно было бы подобрать специфическую химическую модификацию или найти соединение, специфически связывающееся с данными участками и препятствующее их контакту с $Val^{\beta 6}$. Такой способ лечения одной из молекулярных болезней мог бы спасти жизнь многим людям.

Дополнительную информацию о структуре волокна можно получить путем измерения оптического поглощения на ориентированных образцах. Природа этих эффектов, называемых линейным дихроизмом, обсуждается в гл. 7. Гемовые группы поглощают поляризованный свет преимущественно тогда, когда плоскость гема ориентирована параллельно оси поляризации. Поскольку тетрамер содержит четыре гема, непосредственная интерпретация результатов поляризационных измерений довольно трудна. Однако найденные результаты позволяют ограничить возможные ориентации тетрамера в трубке. Некоторые разрешенные ориентации действительно допускают прямое участие $\beta 6$ в межсубъединичном контакте.

Альтернативный подход заключается в вычислении энергетически наиболее благоприятной ориентации тетрамеров в волокне известной структуры. Для простоты предполагается, что каждый тетрамер гемоглобина связан со следующим винтовой осью симметрии. Таким образом, обходясь только одним индивидуальным тетрамером, можно получить все возможные ориентации молекул в трубке. Расчеты, выполненные к настоящему времени, показывают, что среди различных ориентаций существуют наиболее предпочтительные, однако это еще не позволяет выбрать среди них единственную, которой можно однозначно приписать наиболее стабильную структуру.

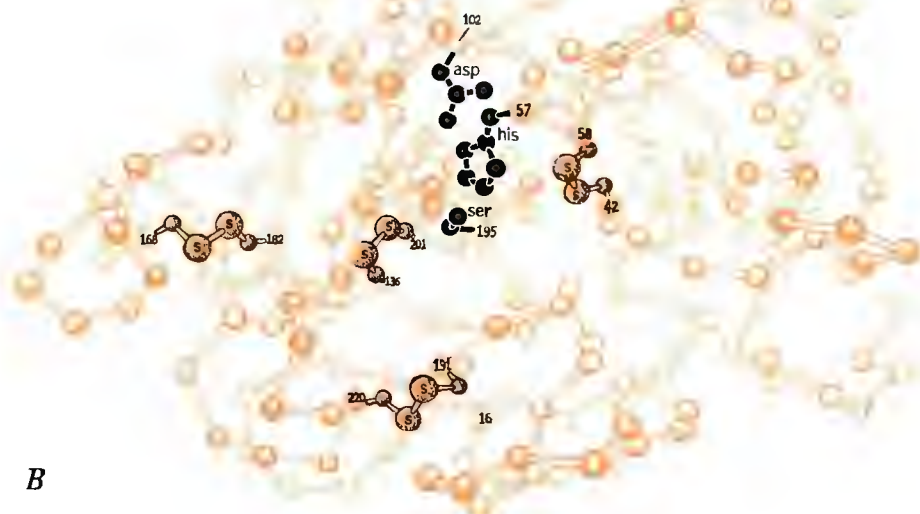
Существуют некоторые особенности физико-химического поведения HbS в растворе. Как правило, кинетика агрегации или ассоциации весьма чувствительна к концентрации. Часто считают, что лимитирующим скорость этапом образования трубки или волокна является ассоциация нескольких субъединиц при формировании зародыша кристаллизации, который затем быстро растет, присоединяя к себе другие структурные единицы (см. рис. 2.21,*В*). Для простых механизмов скорость образования зародыша и, следовательно, скорость агрегации или полимеризации пропорциональна c^n, где c — концентрация свободных субъединиц, а n — число субъединиц в зародыше. Обычно n равно 2 или 3. Однако Хофрихтер, Росс и Итон обнаружили, что для кинетики образования трубок гемоглобина S $n \geqslant 30$ (см. рисунок).

A

РИС. 2.20. Третичная структура трех родственных сериновых протеаз. *А*. Трипсин. *Б*. Химотрипсин. *В*. Эластаза. Здесь показаны только положения α-углеродных атомов, а также несколько боковых групп в области активного центра (черные кружки) и дисульфидные мостики (красные кружки). Все три протеазы расщепляют пептидные связи одинаковым образом, но трипсин преимущественно расщепляет их в участках, расположенных после положительно заряженных остатков, химотрипсин — после больших гидрофобных остатков, а эластаза — после малых гидрофобных остатков. Различную специфичность к субстрату можно понять, анализируя различие в структуре связывающих карманов, примыкающих к каталитическому центру. (Рисунки Ирвинга Гейса.)

102
asp
57
his
ser
195
58
42
168
182
201
136
122
191
220
16
146
149
265

Б


102
asp
57
his
ser
195
58
42
168
182
201
136
191
220
16

В

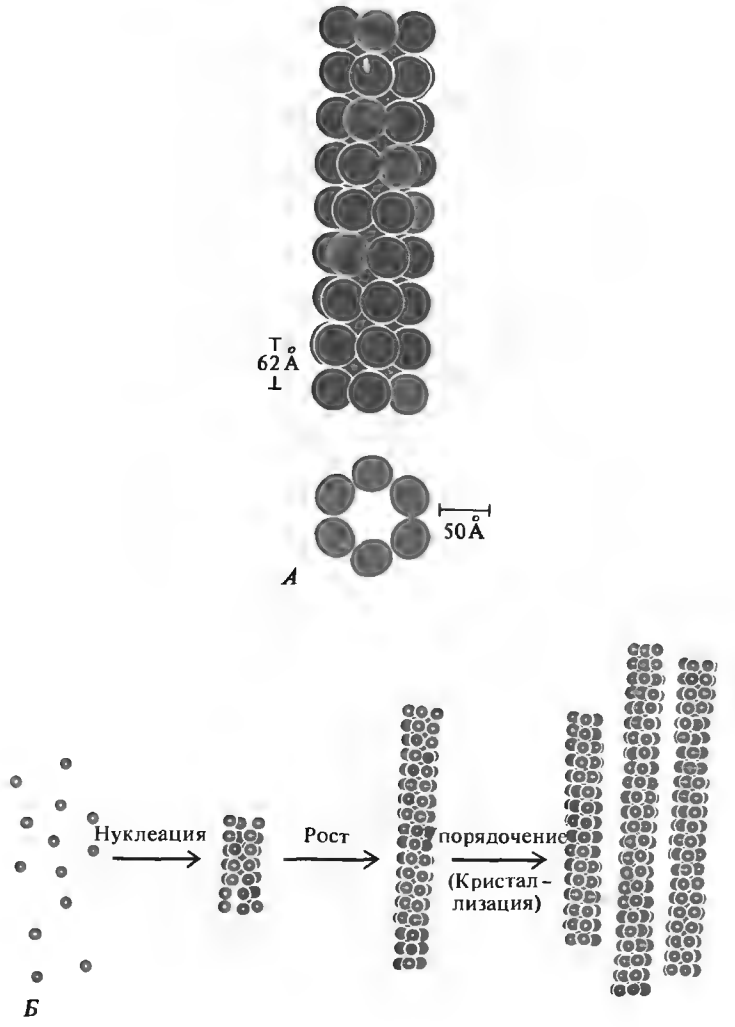
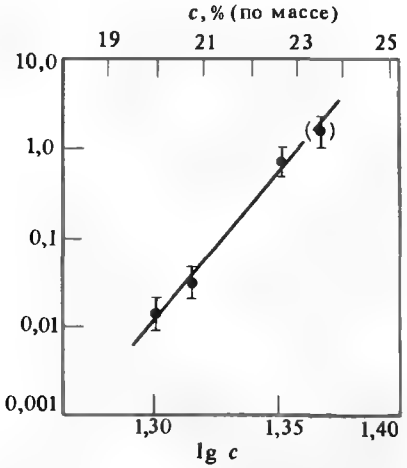


РИС. 2.21. Некоторые особенности агрегации дезоксигемоглобина S, который ответствен за заболевание серповидноклеточной анемией. *А*. Предполагаемая структура отдельного спирального волокна. *Б*. Схематическое изображение механизма формирования агрегата при пониженном давлении кислорода. Нуклеация (образование зародыша кристаллизации) — самый медленный этап, лимитирующий скорость этого процесса. Как только зародыш образуется, дальнейший рост длинных волокон происходит довольно быстро. Волокна затем агрегируют, располагаясь параллельно друг другу, и могут разместиться в эритроците только при соответствующем изменении его формы.



Зависимость временной задержки процесса нуклеации ($1/t$, мин$^{-1}$) от суммарной концентрации гемоглобина S при 30°C. Тангенс угла наклона прямой, построенной в логарифмических координатах, равен 32. Таким образом, скорость нуклеации примерно пропорциональна c^{32}, откуда следует, что наименьшие стабильные центры содержат по крайней мере 32 молекулы гемоглобина. [J. Hofrichter et al., Proc. Natl. Acad. Sci. USA, **71**, 4864, (1974).]

Это поразительный результат. Столь большое значение n никак не вытекает из известной нам структуры трубки. Но следующие отсюда выводы о возможных методах лечения серповидноклеточной анемии очень важны. Скорость образования серповидных клеток существенно зависит от концентрации гемоглобина в эритроците, от объема последнего, от любых термодинамических параметров, контролирующих эти переменные, а также от прочности контактов молекул гемоглобина друг с другом. Если мы узнаем, как уменьшить скорость образования серповидных клеток, то сможем лучше лечить эту болезнь. Эритроцит находится в периферических капиллярах очень недолго. Как только он возвращается опять в легкие и снова окисляется, трубки перестают быть устойчивыми. Таким образом, если скорость образования серповидных клеток достаточно мала, то эритроциты не успевают превратиться в серповидные за время прохождения через капилляры.

2.4. Вторичная структура

Мы уже указывали, что аминокислотные цепи могут образовывать высокоупорядоченные локальные структуры со спиральной симметрией. Эти структуры важны по многим причинам. Спираль ввиду ее симметрии легко распознать и описать. На рис. 2.22 изображен один из участков карты электронной плотности миоглобина, построенной методом рентгеноструктурного анализа. Для непривычного глаза большая часть карты покажется совершенно хаотичной. Однако спирали сразу же бросаются в глаза как темные круги (вид с торца) или как периодически расположенные темные пятна (вид сбоку) (эти

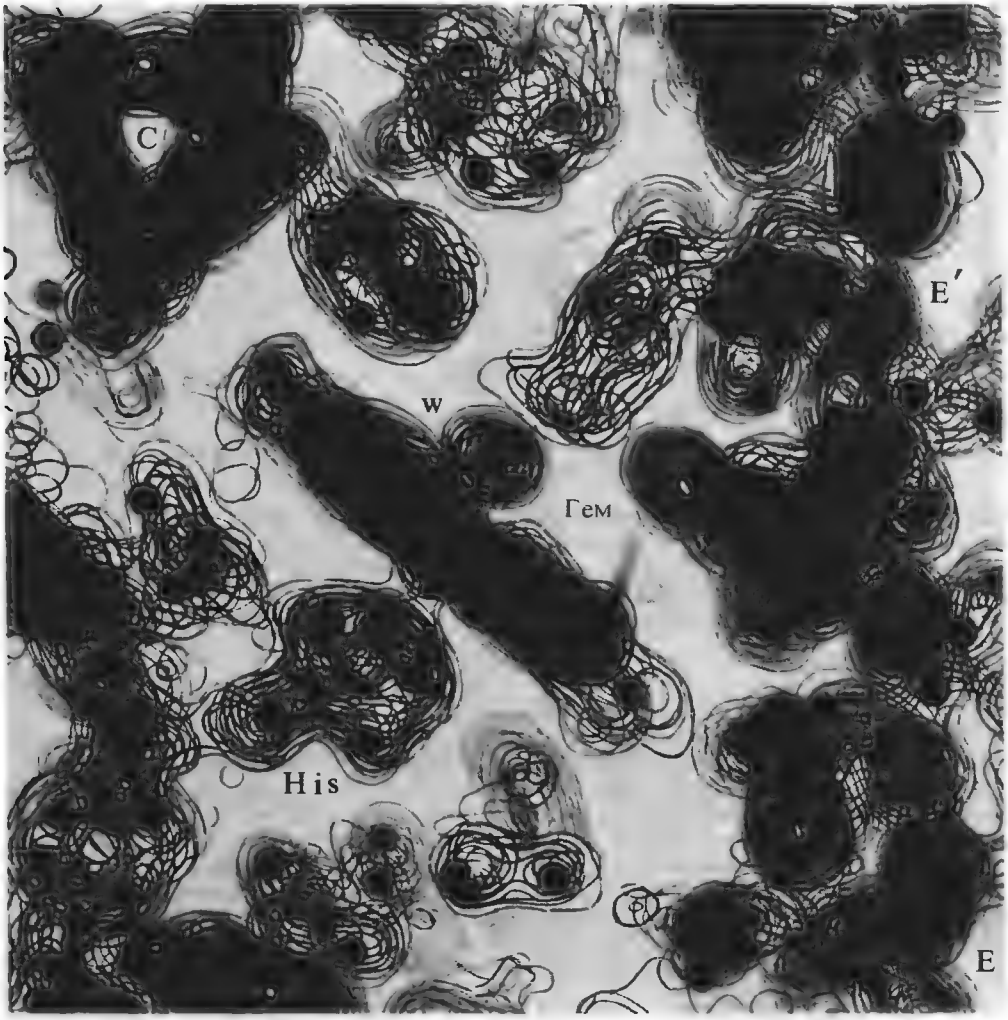


РИС. 2.22. Часть карты электронной плотности для миоглобина, полученной методом рентгеноструктурного анализа. Две α-спирали, расположеиные перпендикулярно плоскости рисунка, обозначены буквой С; одна спираль, видимая сбоку, расположена между *E* и *E′*. Показаны также вид гема сбоку и сязанные с ним остаток гистидина и молекула воды W (ср. с рис. 2.7, *Г*) (М. Ф. Перутц).

области обозначены буквами *C*, *E* и *E′* на рис. 2.22). Квалифицированный кристаллограф различает спиральные участки на картах электронной плотности и при худшем разрешении. Появление спиралей часто служит важным указанием на то, что расшифровка структуры идет в правильном направлении. Точная локализация спиралей дает возможность перейти к определению структуры остальной части молекулы.

Спирали стабилизируются ближними взаимодействиями. Таким образом, в отличие от большей части белка они могут быть выделены (либо с помощью ферментов, либо путем использования модельных соединений) и изучены независимо. Это важный шаг общей стратегии «разделяй и властвуй». В гл. 5 обсуждается вопрос о том, почему определенные спиральные структуры внутренне стабильны, а в гл. 20 — факторы, которые определяют их образование и стабильность.

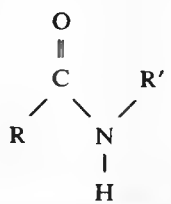
Наличие спиральных структур оказывается чрезвычайно полезным при спектроскопических исследованиях белков. Как правило, спектры чувствительны к локальной структуре отдельных остатков или образуемых ими пар. В сложной третичной структуре существует так много локальных конформаций, что спектр может стать неинформативным. Однако подобие локальной структуры остатков в спиралях или β-слоях гарантирует, что их вклад будет заметен. Именно благодаря этому многие конформационные изменения белка, сопровождающиеся изменениями вторичной структуры, можно охарактеризовать спектроскопически. Это не единственный известный нам уровень структурной организации, но часто именно он оказывается наиболее подходящим для количественного исследования.

α-СПИРАЛЬ

Наиболее известный элемент вторичной структуры полипептидов — правая α-спираль. Она довольно точно воспроизведена на рис. 2.23,*А*. Основные характеристики спирали: на один виток приходится 3,6 аминокислотных остатка; на один остаток спирали приходится 1,5 Å, шаг спирали $3{,}6 \cdot 1{,}5 = 5{,}4$ Å. Диаметр спирали без учета боковых групп составляет примерно 6 Å. Каждый карбонил пептидной группы (без учета концевых эффектов) в α-спирали служит акцептором при образовании водородной связи для донора N—H пептидной группы, отстоящей от него на расстоянии четырех остатков. Таким образом, если расположить α-спираль N-концом вверх, а C-концом вниз (как изображено на рис. 2.23,*А*), то все C—O-группы будут направлены вверх, а все N—H-группы — вниз.

Боковые группы направлены в сторону от спирали, поскольку все остатки являются L-аминокислотами (см. также рис. 1.1). Таким образом, в эту структуру могут встраиваться любые остатки с гибкими боковыми цепями при условии, что они не столь объемны, чтобы создавать стерические препятствия. Одако, как будет показано в гл. 5, пролин по стерическим соображениям не укладывается в α-спираль. Конечно, для некоторых боковых цепей нахождение внутри α-спирали энергетически невыгодно. Так, например, поли-L-глутамат при высоких значениях pH будет полианионом. Электростатическое отталкивание между боковыми группами сильно дестабилизирует спираль и при низкой ионной силе способствует переходу к вытянутой структуре, поскольку это ведет к максимизации расстояний между отрицательно заряженными боковыми группами и тем самым к минимизации энергии электростатического отталкивания.

Одно из наиболее впечатляющих достижений структурной химии — предсказание Полингом и Кори существования регулярной структуры типа α-спирали (а также существования правильных структур в нескольких других упорядочённых полипептидных конформациях) за 6 лет до того, как она впервые была обнаружена в кристаллической структуре миоглобина. Интересно проследить за рассуждениями, которые привели авторов к такому выводу. Длины связей и углы были известны из рентгеноструктурных исследований целого ряда малых пептидов. Разумно было предположить, что их значения не изменятся в любой вторичной структуре. Практически во всех известных структурах размеры пептидных групп одинаковы, а входящие в них группы C—O и N—H находятся в плоской *транс*-конформации:



Предположение о том, что такая конформация сохраняется всегда и во всех остатках, очень упростило дальнейшее рассмотрение и сразу же навело на мысль: если каждый пеп-

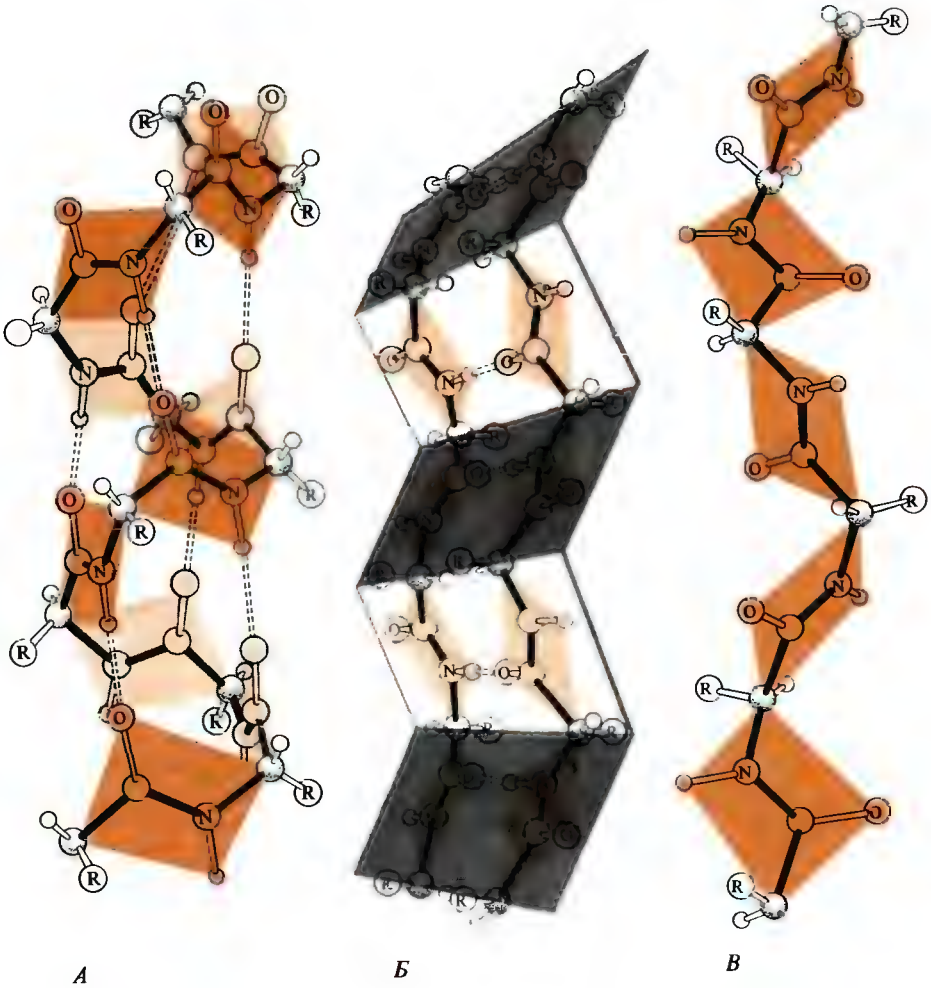


РИС. 2.23. Три наиболее общих типа вторичной структуры полипептидов. Изображены только атомы скелета полипептидной цепи. Параллелограмм показывает плоскость, в которой расположена пептидная связь. *А*. Правая α-спираль. *Б*. Две антипараллельные цепи в β-слое. Обратите внимание, что чередующиеся боковые группы располагаются на противоположных сторонах β-слоя. *В*. Левая спираль полипролина II. (Рисунки Ирвинга Гейса.)

тидный скелет имеет одну и ту же структуру, то и весь полипептидный скелет должен обладать регулярной вторичной структурой типа какой-либо спирали. Исходя из этого, Полинг и Кори стали искать такие полипептидные спирали, которые можно было бы построить из обычных аминокислотных остатков. Чтобы ограничить количество возможных гипотетических моделей, они наложили два дополнительных требования: 1) количество водородных связей между пептидными группами N—H · · · O = C должно быть максимальным и 2) сами связи должны быть по возможности близки к линейным. Эти требования основывались на том, что энтальпия водородной связи для системы, находящей-

ся в газообразном состоянии, составляет 5 — 7 ккал/моль, тогда как конформационные энергии поворота вокруг одинарных связей значительно меньше и, следовательно, можно заключить, что энергии водородных связей играют наиболее важную роль.

β-СЛОИ И ДРУГИЕ ТИПЫ ВТОРИЧНОЙ СТРУКТУРЫ

Из рис. 2.23 следует, что правая α-спираль прекрасно удовлетворяет критериям Полинга и Кори. Было также постулировано существование левой спирали и двух типов β-слоев. Антипараллельный β-слой изображен на рис. 2.23,*Б*. Каждая отдельная цепь является спиралью с осью симметрии второго порядка, т.е. образует складчатую структуру. Все пептиды принимают участие в образовании водородных связей, однако все эти связи образуются между различными цепями. Геометрия пептидного скелета в β-слое соответствует наиболее вытянутой конформации цепи, еще допускаемой при нормальных значениях длин связей и валентных углов; смещение вдоль оси равно 3,47 Å на остаток. Во многих белках объединение параллельных и антипараллельных β-цепей приводит к образованию большого структурного остова (рис. 2.24, а также Дополнение 2.3).

Все четыре вторичные структуры, обнаруженные либо в белках, либо в модельных полипептидных системах, совпадали с предсказанными Полингом и Кори. Однако, как будет показано в гл. 5, критерий, первоначально используемый для их построения, оказался не совсем точным и был изменен с учетом современного представления о силах, стабилизирующих конформации макромолекул. Часть трудностей касались эффективной прочности водородных связей. Другое серьезное затруднение возникло в связи с влиянием боковых цепей. Именно они в основном определяют, образует ли данная аминокислотная последовательность неупорядоченную структуру, α-спираль, β-слой или некоторые другие структуры, обнаруженные в белках. Спираль 3_{10} иногда находят в коротких участках структуры глобулярных белков. Это, по существу, вариант α-спирали, в котором на один виток приходится 3 остатка вместо 3,6. (В коротких участках структуры глобулярных белков иногда находят и два других варианта спиралей: α_{II}-спираль и π-спираль. Описание этих структур можно найти в книге Дикерсона и Гейса (1969).) Гораздо шире распространена β-структура, изображенная на рис. 2.25. Она дает пептидной цепи возможность круто поворачивать, сохраняя при этом энергетически выгодные значения конформационных углов и одну внутрицепочечную водородную связь. β-Изгибы справедливо называют элементами вторичной структуры, поскольку они возникают в результате ближних взаимодействий и их можно рассматривать как начало спирали с нулевым шагом.

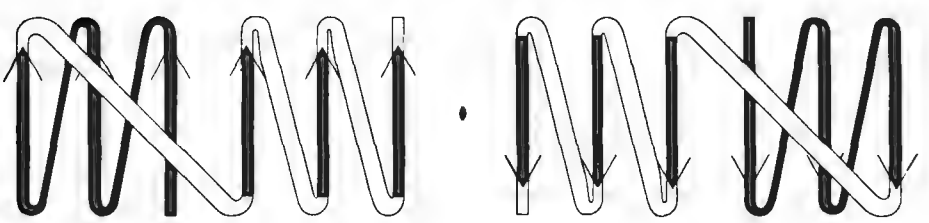


РИС. 2.24. Схематическое представление системы водородных связей в β-слое, образованном двумя одинаковыми субъединицами алкогольдегидрогеназы из печени лошади. В каждой субъединице существует шестицепочечный параллельный β-слой. Между двумя субъединицами располагается двухцепочечный антипараллельный слой. Указано положение оси симметрии C_2. [M.G.Rossman et al. In: The Enzymes, 3d ed., vol. 9, ed. P.Boyer (New York: Academic Press, 1975), p.61.]

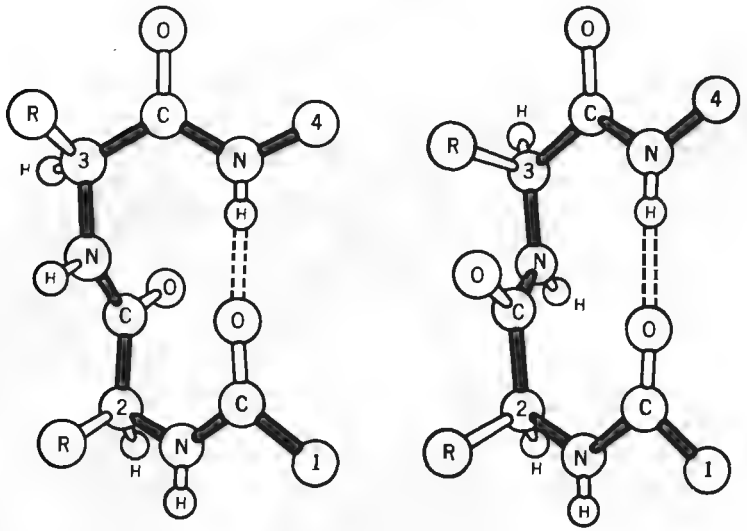


РИС. 2.25. Структура β-изгибов двух типов. (Рисунок Ирвинга Гейса.)

ПОЛИПРОЛИНОВЫЕ СПИРАЛИ И КОЛЛАГЕН

Последний из распространенных типов вторичной структуры — полипролиновая спираль. Пролин не может образовывать обычных α- и β-структур из-за ограничений, налагаемых на его скелет пятичленным кольцом. Однако он может образовывать две односпиральные структуры, которые являются уникальными среди полипептидных вторичных структур, так как не содержат водородных связей. Полипролин I — левая спираль, содержащая 10 остатков на 3 витка. Все пептидные связи в ней находятся в *цис*-конформации, что довольно редко встречается в обычных пептидных структурах. Полипролин II — левая спираль, и ее пептидные связи находятся в *транс*-положении. Взаимопревращение этих двух форм может быть вызвано изменениями в растворителе. В нормальных водных буферах стабильной формой является спираль II. Она имеет 3 остатка на виток и смещение вдоль оси спирали 3,12 Å на остаток по сравнению с 1,5 Å в α-спирали. Структура полипролина II, изображенная на рис. 2.23, представляет собой более вытянутую цепь, чем α-спираль. В такой вытянутой спирали каждая боковая цепочка может быть сильно удалена от всех остальных. На самом деле полипролиновые спирали в значительной мере стабилизированы именно в результате таких стерических ограничений.

Поли-L-пролин — синтетическая структура. Цепь из одних пролинов или даже участки с большим количеством пролинов в большинстве белков встречаются крайне редко, и поэтому очень трудно найти вторичную структуру, похожую на спираль поли-L-пролина II. Удивительное исключение составляют, однако, белки семейства коллагена. Они имеют в высшей степени необычный состав, в котором преобладают пролин, гидроксипролин и глицин. Тропоколлаген содержит три полипептидные цепи: две тождественные, называемые α_1-цепями, и еще одну, также близкую к ним, — α_2-цепь. Мол. масса каждой около 95 000. Последовательности всех цепей имеют в каждом третьем положении глицин, тогда как одно (или оба) из промежуточных положений занимает пролин или гидроксипролин (см. рис. 2.15). Высокое содержание пролина в коллагеновых цепях ограничивает возможные структуры полипролиновыми спиралями. Изолированный тропоколлаген имеет мол. массу 285 000 и представляет собой палочкообразную молекулу дли-

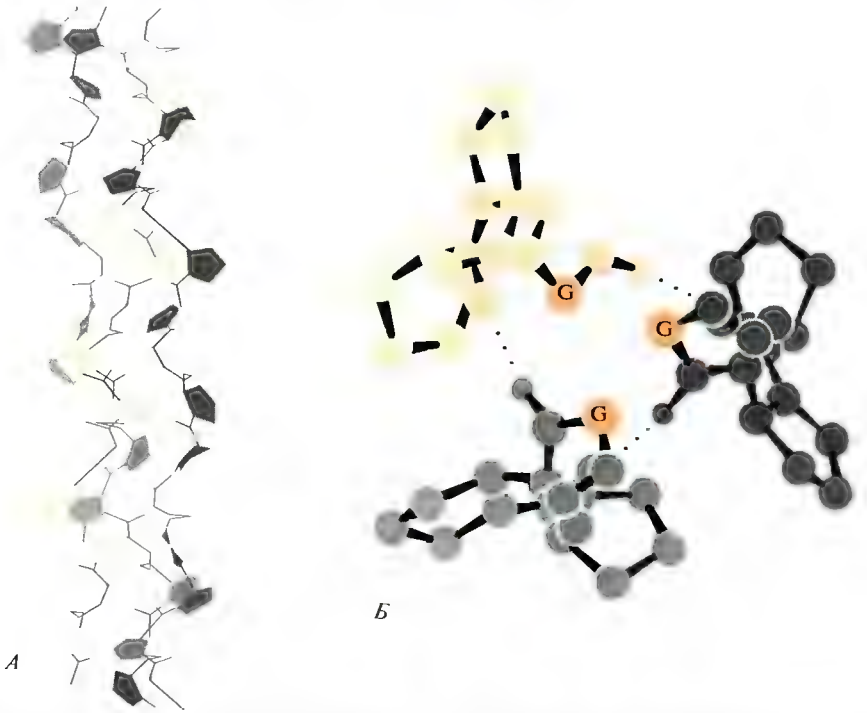


РИС. 2.26. Трехцепочечная спираль коллагена. *А*. Вид на структуру перпендикулярно оси спирали. *Б*. Вид вдоль оси спирали: показаны межцепочечные водородные связи. Изображена последовательность $(—Gly—Pro—Pro—)_n$. Отметим, что остатки пролина находятся снаружи, а глицина — внутри спирали. Каждым третьим остатком должен быть глицин, так как в центре спирали нет места для большего остатка. Все C^{α}-атомы глициновых остатков цветные. [L.Stryer, Biochemistry (San Francisco: W.N.Freeman and Company, 1975).]

ной 3000 Å и только 15 Å в диаметре. Он образует тройную спираль из трех закрученных относительно друг друга спиралей (рис. 2.26). Эта тройная спираль содержит межцепочечные водородные связи. Все пролиновые кольца выставлены наружу, все глицины расположены внутри, так как основная особенность последних — малый объем, а три цепи упакованы настолько плотно, что для любого остатка с бóльшим объемом не хватило бы места. Глицины расположены таким образом, чтобы на каждом уровне тройной спирали (расстояние по вертикали между остатками равно 2,9 Å) в одной из цепей всегда находился глицин.

Псевдоповторяющаяся последовательность отдельных цепей тропоколлагена создает проблему для формирования тройной спирали. Три цепи могут быть сдвинуты друг относительно друга на три остатка или на любое число остатков, кратное трем, и вместо палочкообразной коллаген будет иметь структуру с «разлохмаченными» концами, что приведет к беспорядочному межмолекулярному агрегированию. Между тем биологическая роль коллагена состоит в образовании строго ориентированных волокон. Эти жесткие, очень длинные волокна придают механическую прочность соединительной ткани. В волокне молекулы тропоколлагена организованы в регулярные четвертичные структуры (рис. 2.27). «Разлохмаченные» концы препятствовали бы правильной укладке. Поэтому для обеспечения правильного расположения цепей тропоколлагена друг относительно друга синтезируется предшественник — проколлаген. Молекулярная масса каждой цепи проколлагена равна 120 000. Небольшие дополнительные участки молекулы, расположенные на одном ее конце, имеют свойства, более характерные для обычных глобулярных белков. Они сворачиваются, сшиваются дисульфидными связями и образуют структурное

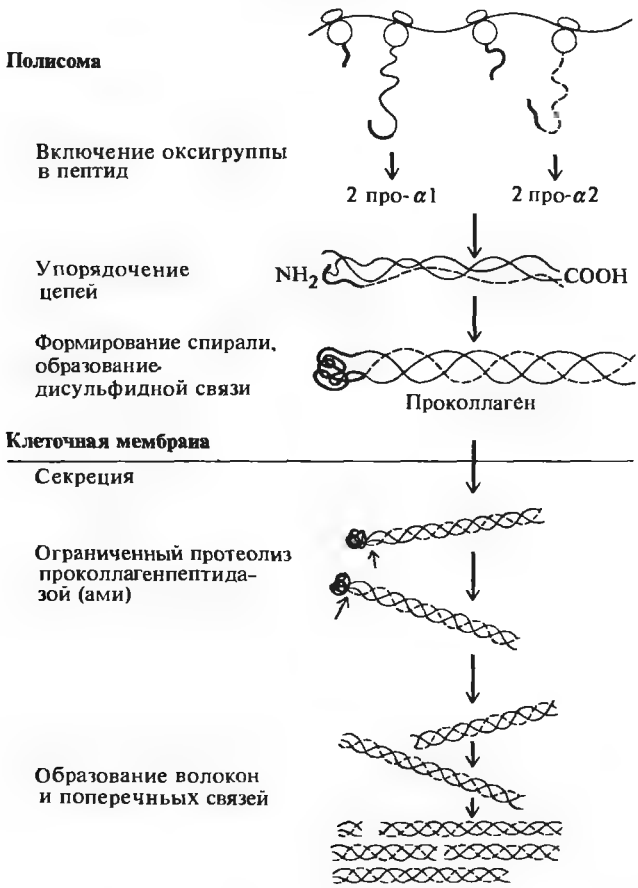


РИС. 2.27. Некоторые этапы биосинтеза коллагена и образования коллагеновых волокон. [P.Bornstein, Ann.Rev. Biochem, **43**, 567 (1974).]

ядро, которое обеспечивает точную укладку цепей в коллагеновой спирали. Затем эти дополнительные участки удаляются в результате протеолиза.

РАСПРОСТРАНЕННОСТЬ ВТОРИЧНЫХ СТРУКТУР В БЕЛКАХ

В типичном белке, трехмерная структура которого определена методом рентгеноструктурного анализа, обычно около 60% остатков так или иначе участвуют в формировании вторичных структур. В табл. 2.7 приведены имеющиеся сегодня данные. Из нее видно, что большинство белков содержит в значительных количествах как α-, так и β-структуры. Ферредоксин *P. aerogenese* представляет редкое исключение, поскольку он фактически не содержит поддающейся обнаружению вторичной структуры. Это легко объяснить. Этот белок состоит всего лишь из 54 аминокислот. Восемь из них — Cys (см. рис. 2.15). Они группируются по четыре с образованием лигандов для двух железосерных кластеров кубической структуры (см. рис. 2.7).

Схематическое изображение третичной структуры ферредоксина приведено на рис. 2.28. Как мы видим, пептидная цепь едва может обеспечить связь одного центра координации с другим. Эти ограничения не дают возможности образоваться вторичным структурам. Такие трудности возникают не у всех белков, содержащих простетические группы; так, например, гемоглобин и миоглобин содержат очень много α-спиралей, однако они свернуты таким образом, что образуется место связывания гема (см. рис. 1.4 и 2.18).

Таблица 2.7

ВТОРИЧНЫЕ СТРУКТУРЫ В БЕЛКОВЫХ КРИСТАЛЛАХ[1)]

Белок	Спирали			β-слои				β-изгибы	
	%	число	типы структуры	%	число β-слоев	число цепей	типы структуры	число β(I)	число β(II)
Карбоангидраза	20	7	α, α_{II}, 3_{10}	37	2	13	п, а, (ц)	3	1
Карбоксипептидаза А	38	9	α, α_{II}	17	1	8	п, а	10	3
Химотрипсин (эластаза, трипсин)	14	3	3_{10}, α	45	2	12	а, 2ц	11	6
Конканавалин А	2	1	—	57	3	18	а, а2а	—	—
Цитохром b_5	52	6	α, 3_{10}	25	1	5	п, а	1	1
Цитохром *c*	39	5	—	—	—	—	—	3	3
Ферредоксин	—	—	—	4	1	2	а	—	—
Флаводоксин	30	4	—	30	1	5	—	—	—
Инсулин	52	3	α, 3_{10}	6	1	1	а, а2а	—	—
Лактатдегидрогеназа (или малатдегидрогеназа)	45	10	—	20	3	12	п, а	6	4
Лизоцим	40	6	α, α_{II}, 3_{10}	12	2	6	а	4	2
Миоген	57	6	α, 3_{10}	4	1	2	—	5	—
Миоглобин (гемоглобин)	79	8	α, 3_{10}, π	—	—	—	—	6	—
Нуклеаза	24	3	—	14	1	3	а	—	—
Ингибитор трипсина из поджелудочной железы	28	2	α, пп	33	1	3	а	—	—
Папаин	28	5	—	15	1	7	а	—	—
Рибонуклеаза	26	3	α, 3_{10}	35	1	6	(п), а	1	1
Субтилизин	31	8	—	16	2	8	п, а	16	1
Термолизин	38	7	—	22	1	10	п, а, (ц), кл	—	—

[1)] Проценты указывают доли всех остатков, включенных в спирали и β-слои соответственно. Используемые символы означают следующее: α — α-спираль; α_{II} — α_{II}-спираль; 3_{10} — 3_{10}-спираль; π — π-спираль; а — антипараллельная β-структура; а2а — аитипараллельиая β-структура, связанная осью второго порядка; ц — цилиндрическая β-структура; кл — β-структура типа клеверного листа; п — параллельная β-структура; пп — полипролиновая структура; заключение символа в скобки указывает на то, что соответствующая структура искажена. [A. Lilias, M. Rossmann, Ann. Rev. Biochem., **43**, 480 (1974).]

При рассмотрении табл. 2.7 надо учитывать, что в ней приведена специфическая подборка белков. Большинство из них имеет малую молекулярную массу, и их удалось выделить в кристаллической форме в виде хорошо упорядоченных кристаллов, что позволило получить дифракционные данные по нативному белку и его изоморфным производным с высоким разрешением. По всей видимости, значительная доля вторичных структур в молекуле белка способствует образованию упорядоченных кристаллов. Практически наличие спиралей и β-слоев в белках можно установить по результатам спектроскопических исследований растворов белков. Но для того чтобы пойти дальше и выяснить, сколько в структуре спиралей, какой они длины, сколько цепей содержится в β-слое и т.п., обычно

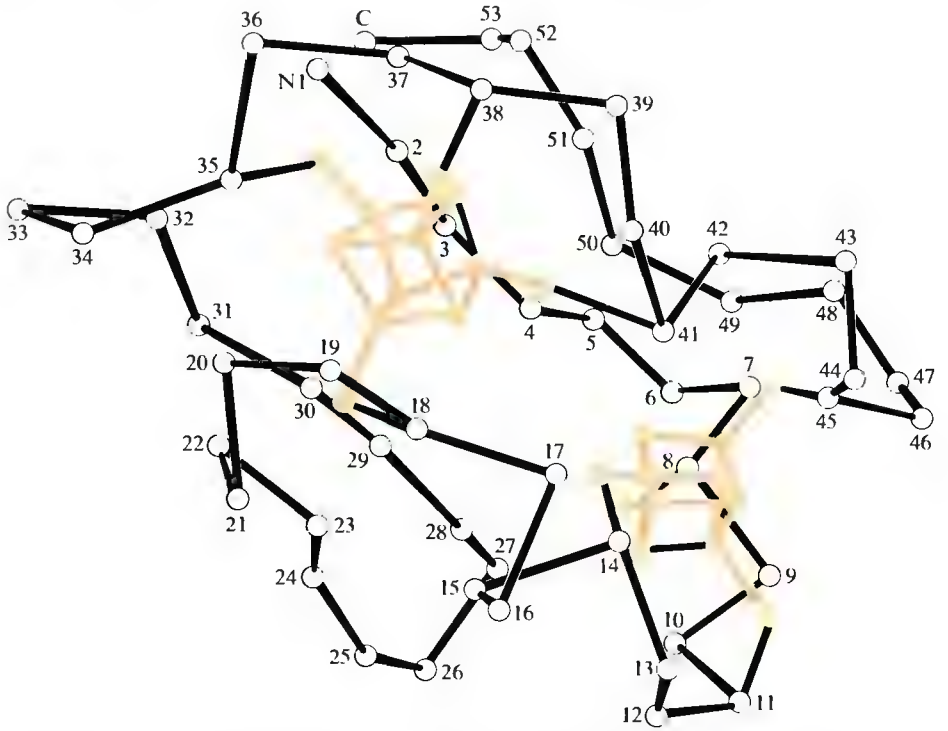


РИС. 2.28. Третичная структура ферредоксина из *P. aerogenes*, определенная методом рентгеноструктурного анализа. Светлые кружки — C^{α}, цветные кружки — Fe, лабильная сера и атом серы в остатке цистеина. Железосерные кубические кластеры изображены цветными. [L.H.Jensen, Ann.Rev. Biochem.,, **43**, 461 (1974).]

надо иметь кристаллическую структуру. Особенности расположения вторичных структур являются одной из основных характеристик третичной структуры, к обсуждению которой мы переходим.

2.5. Третичная структура

Ранее был показан ряд третичных структур. Они очень сложны, и обычно до того, как удается выяснить основные их особенности, приходится немало потрудиться. На одном рисунке очень трудно показать различные аспекты структуры. Некоторые из них можно представить себе только на реальной модели. Объемная модель часто оказывается недостаточно наглядной. Рис. 2.29 изображает такую модель для рибонуклеазы. Из модели видно, что белки — довольно плотные структуры; очевиден также действительно сложный характер поверхности молекулы и формы области связывания (кармана) субстрата. Детали хода полипептидной цепи и характер вторичных структур по такой модели установить невозможно.

Если при построении модели не принимать во внимание водородные атомы, то на ней можно увидеть больше деталей. Такая модель представлена для гексокиназы А на рис. 2.30. Рисунок выполнен с помощью ЭВМ.

На рис. 2.31 — 2.33 изображены более абстрактные модели структуры лизоцима. На них каждой химической связи в скелете и боковых цепях соответствует отрезок прямой

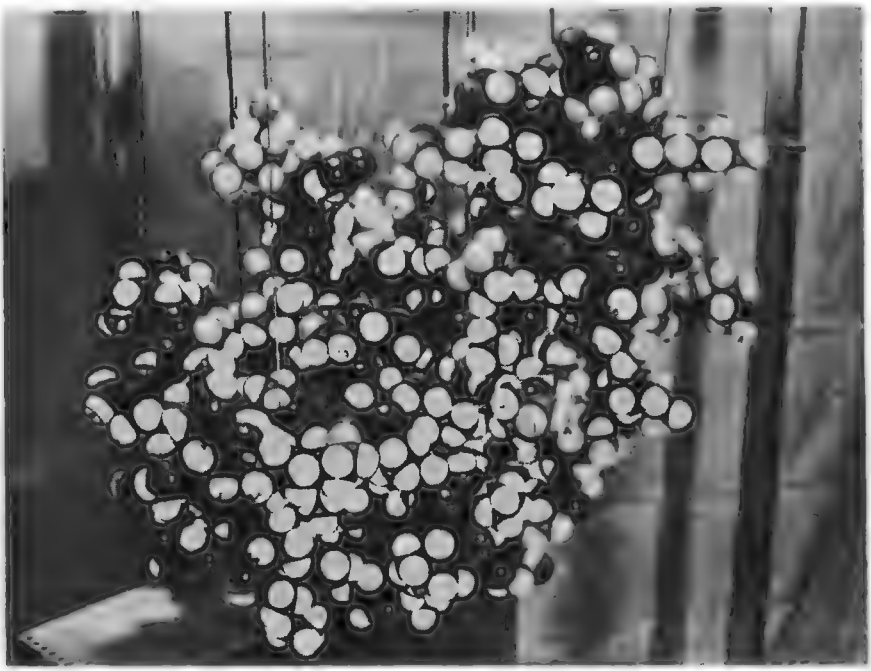


РИС. 2.29. Объемная модель рибонуклеазы S поджелудочной железы быка. Широкий желоб, проходящий вертикально и особенно хорошо заметный в верхней части молекулы, является областью связывания субстрата (ср. рис. 16.14). (Фото предоставлено Ф. М. Ричардсом.)

линии. Художник несколько исказил структуру, чтобы сделать ее более наглядной, но даже в этом случае трудно разобраться в такой запутанной сетке связей. А ведь это еще очень маленький белок. Для больших белков подобные изображения не очень полезны.

Еще более абстрактными выглядят рисунки, на которых изображены лишь C^{α}-атомы и указаны несколько наиболее важных боковых цепей, простетические группы или положения дисульфидных связей. Подобный пример приведен на рис. 1.12, где изображена часть молекулы лизоцима. В качестве другого примера укажем термолизин (рис. 2.34). Большая наглядность моделей, состоящих только из C^{α}-атомов, не должна вводить в заблуждение своей простотой, поскольку реальные структуры значительно сложнее, что легко увидеть из сравнения с другими моделями. И наконец, последний способ изображения третичной структуры иллюстрирует рис. 2.35, где представлена структура карбоангидразы человека. На рисунке особое внимание уделено наиболее характерным особенностям фермента (таким, как скрученный β-слой), тогда как другие элементы структуры показаны в упрощенном виде. На самом деле такой рисунок фактически является лишь наброском структуры. В этой книге мы используем простейшие типы моделей, необходимые для иллюстрации некоторых изучаемых проблем. Следует, однако, помнить, что части структуры, не показанные на рисунке, могут быть для нее очень важны.

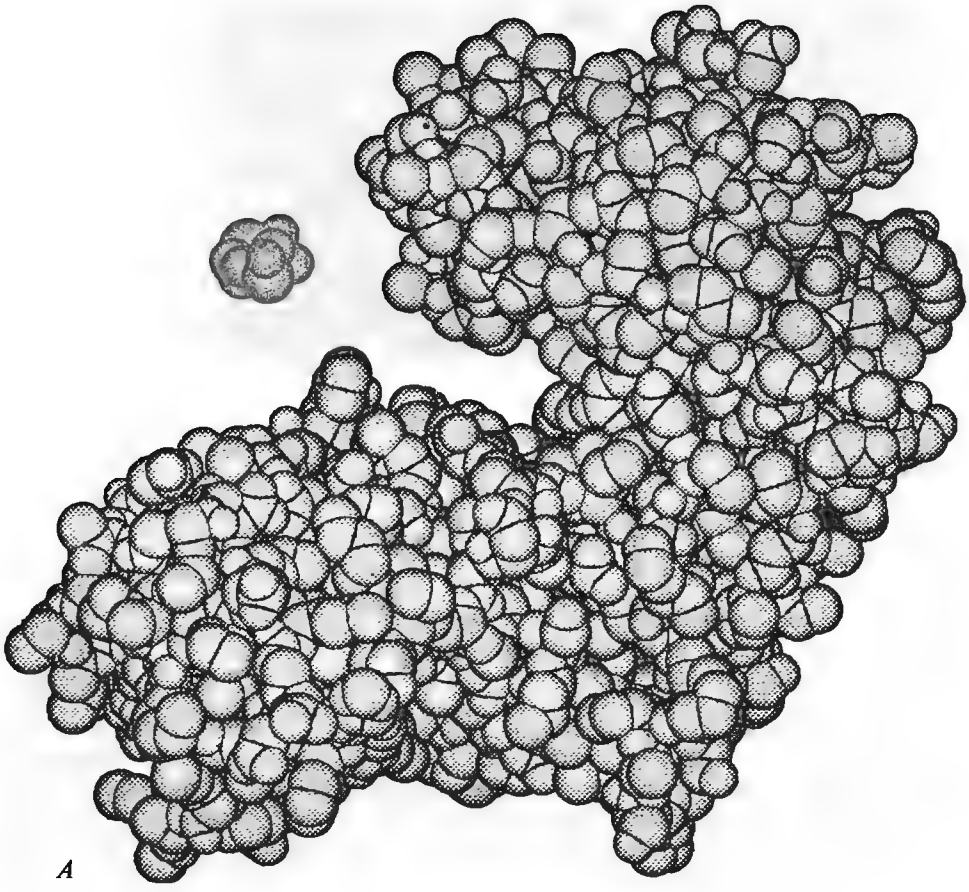
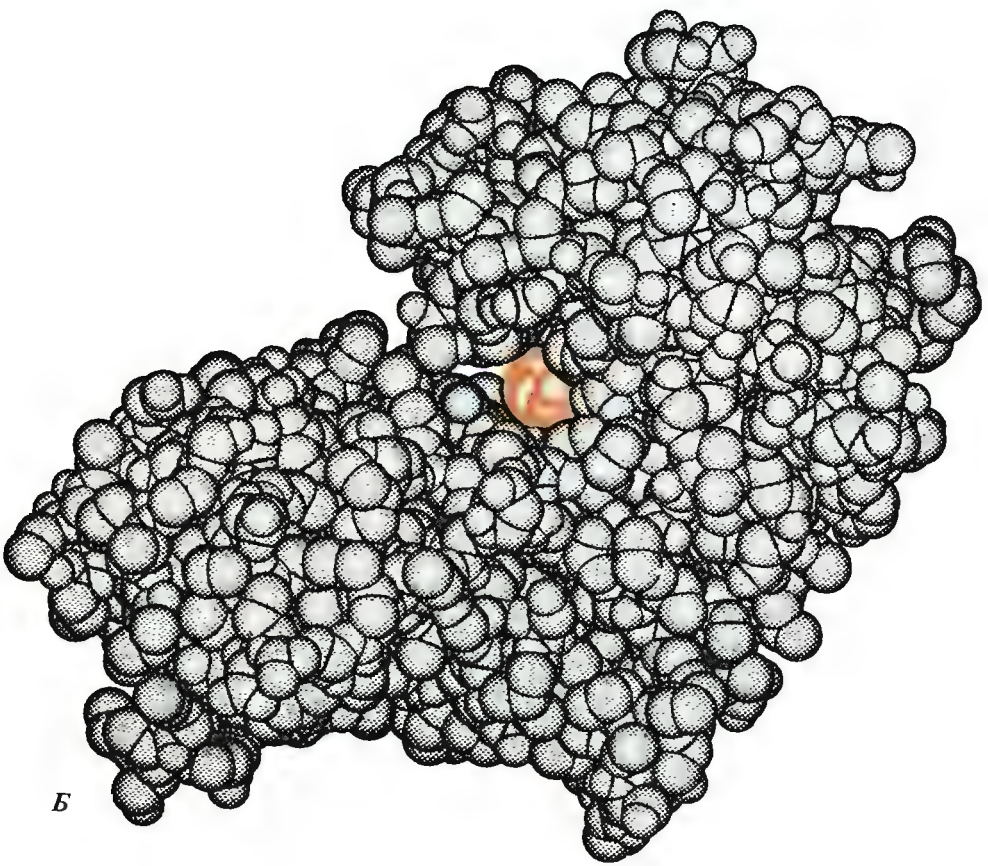


РИС. 2.30. Третичная структура гексокиназы А из дрожжей, нарисованная на электростатическом графопостроителе с помощью ЭВМ. Показаны все атомы, кроме атомов водорода. Последние учитывали путем увеличения радиуса более тяжелых атомов, к которым они присоединены. *А*. Нативная гексокиназа и ее субстрат — глюкоза (цветная молекула). *Б*. Комплекс гексокиназы с глюкозой. Обратите внимание на изменение структуры фермента, вызванное связыванием глюкозы. Фермент так окружает молекулу глюкозы, что для освобождения связанного лиганда необходимы конформационные изменения структуры фермента. (Рисунок предоставлен Т. Стейцем.)



Б

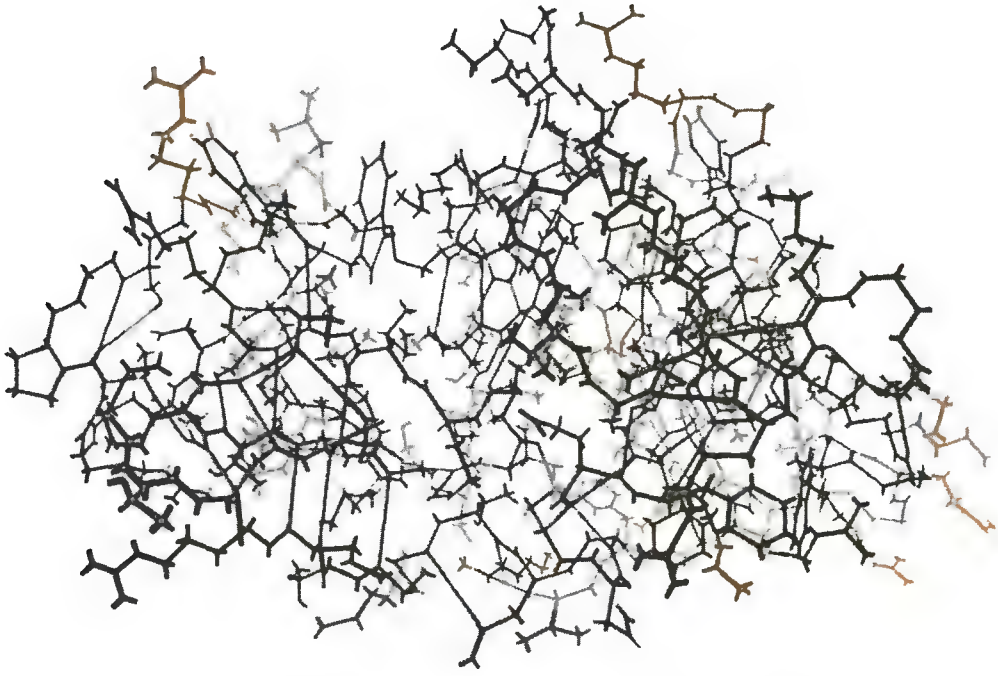
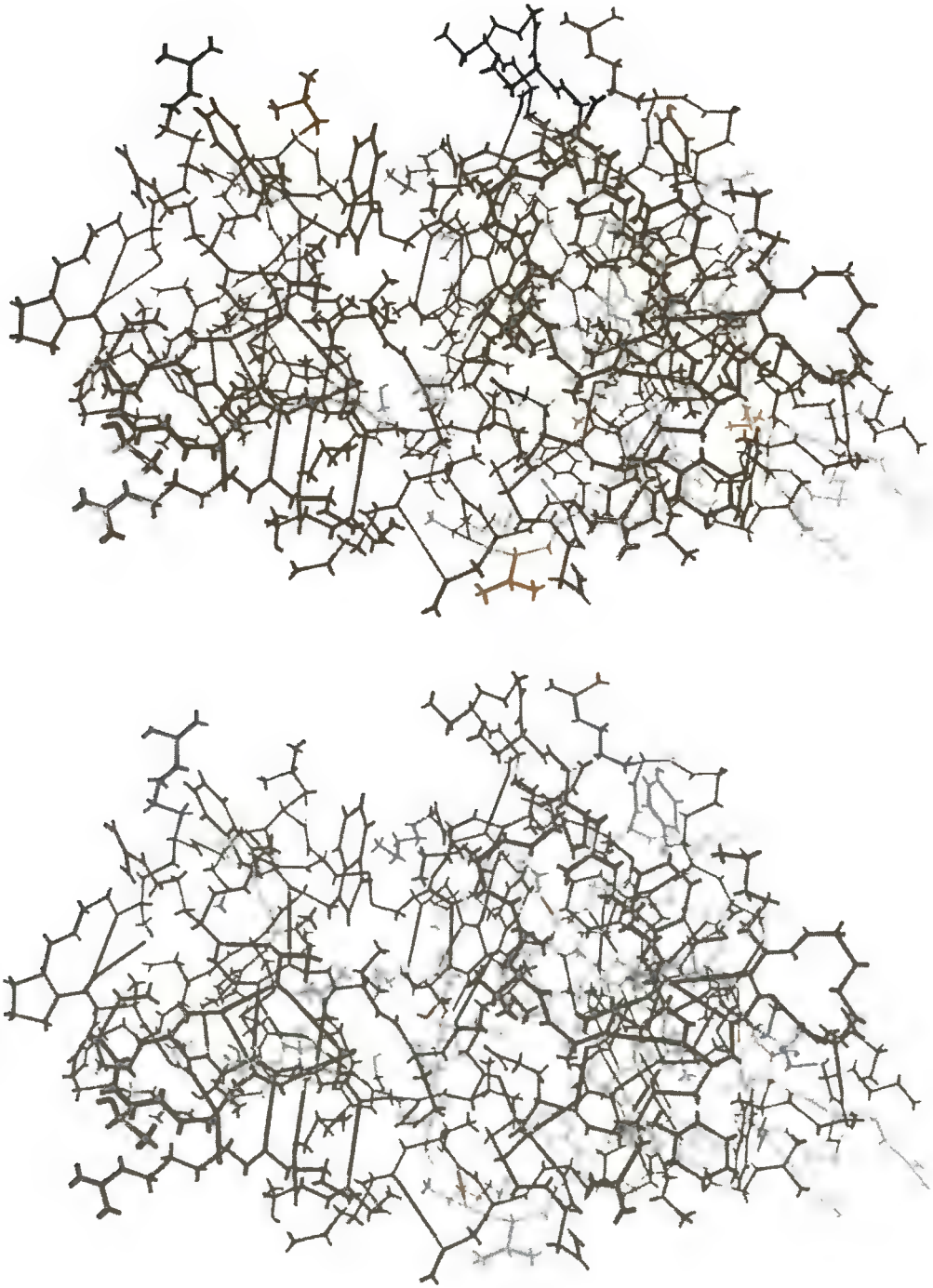


РИС. 2.31. Положение заряженных групп в третичной структуре лизоцима из белка куриного яйца. На этом и на двух следующих рисунках изображены все атомы в молекуле белка. Для наглядности художник несколько исказил структуру. Боковые остатки Lys, Arg, Asp и Glu показаны цветными, так же как и N- и С-концы. (Рисунок Ирвинга Гейса.)

РИС. 2.32. Положение гидрофобных боковых групп в структуре лизоцима из белка куриного яйца. Цветными изображены следующие боковые группы: Trp, Tyr, Phe, Pro, Ile, Leu, Val, Ala, Cys—S—S—Cys и Met (см. подпись к рис. 2.31). Обратите внимание на большой кластер неполярных остатков с правой стороны молекулы и на меньший кластер слева от области связывания сахара. (Рисунок Ирвинга Гейса.)

РИС. 2.33. Положение водородных связей между остатками в структуре лизоцима из белка куриного яйца. Все эти водородные связи показаны цветными. Некоторые связи участвуют в образовании областей вторичной структуры, как β-слой (внизу слева) и две α-спирали, проходящие вертикально справа недалеко от центра молекулы. Другие связи не участвуют в образовании каких-либо наблюдаемых вторичных структур. (Рисунок Ирвинга Гейса.)



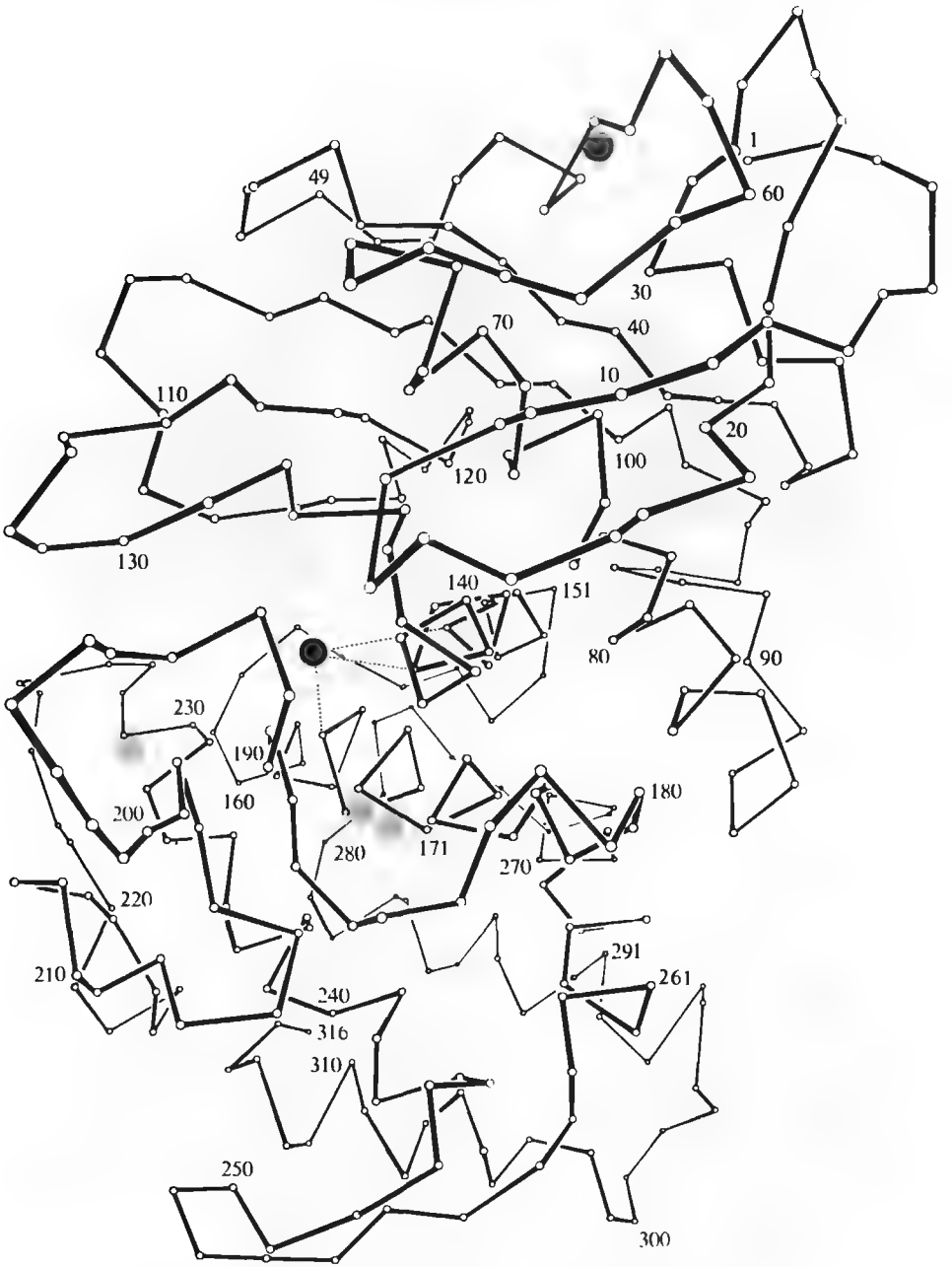


РИС. 2.34. Третичная структура термолизина — протеазы, обладающей чрезвычайно высокой термостабильностью. Обратите внимание, что структура состоит из двух отдаленных друг от друга и независимо складывающихся доменов. C^{α}-атомы изображены светлыми кружками, соединенными друг с другом таким образом, чтобы можно было различить полипептидный скелет. Атом цинка показан зачерненным кружком, четыре иона Ca^{2+} — цветными кружками. Видны многочисленные α-спирали. [B. W. Mattews et al., Nature New Biology, **238**, 37 (1972).]

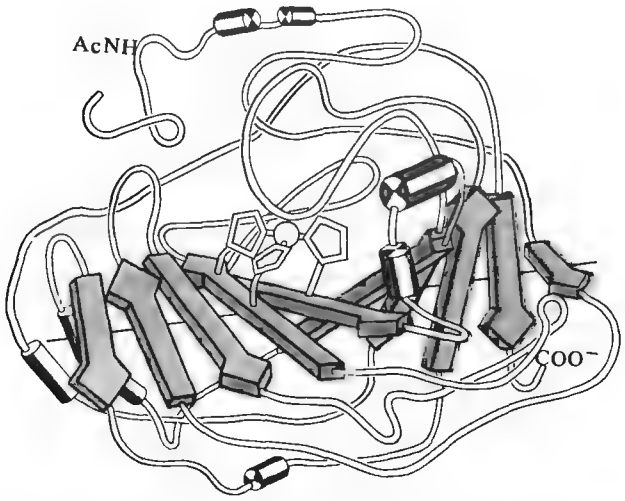


РИС. 2.35. Схематическое изображение третичной структуры карбоангидразы человека, полученное методом рентгеноструктурного анализа. Показаны полипептидная цепь и три остатка гистидина, соединенные координационными связями с ионом цинка в активном центре молекулы. Спирали изображены в виде цилиндров; отдельные цепи β-слоев — в виде стрелок, направленных от N-конца к С-концу. Обратите внимание на скрученность β-слоев [K.K. Kannan et al., Cold Spring Harbor Symp. Quant. Biol., **36**, 221 (1971).]

ОБЩАЯ ОРГАНИЗАЦИЯ ПЕПТИДНОГО СКЕЛЕТА

Проще всего на простейших моделях выявляются вторичные структуры. В табл. 2.7 суммированы данные о вторичной структуре ряда белков. Эти белки содержат от 100 до 300 аминокислотных остатков на субъединицу. Данные о доле остатков в α-спиральных областях и числе α-спиралей, приведенные в таблице, указывают на то, что типичная α-спираль в глобулярном белке состоит примерно из 10 остатков. Длина такой спирали составляет приблизительно 15 Å, что примерно равно радиусу субъединицы в этих белках. Число β-слоев намного меньше, но каждый содержит значительно больше остатков. Большинство обширных β-слоев состоит из 4-6 цепей и содержит от 20 до 40 остатков. Они могут занимать значительные области в третичной структуре, но идентифицировать их труднее, чем α-спирали. Антипараллельные β-слои встречаются несколько чаще, чем параллельные, но многие из слоев содержат как те, так и другие. В качестве примера можно привести расположение элементов вторичной структуры в молекуле карбоангидразы человека (рис. 2.35). Отметим, что обширный β-слой перекручивается, чем-то напоминая по форме левый пропеллер. Такое перекручивание характерно для всех обширных β-слоев в глобулярных белках. Это можно понять, если учесть наиболее выгодные углы внутреннего вращения для отдельных остатков β-слоя. Более подробно β-структура описана в Дополнении 2.3.

В некоторых белках два таких перекрученных слоя формируют структуру, похожую на бочонок (см., например, рис. 2.43). α- и β-структуры не имеют ярко выраженной тенденции располагаться в каких-то определенных местах третичной структуры, например внутри или на поверхности, в области N- или С-конца и т.д. Не существует также какой-либо тенденции элементов вторичной структуры образовывать кластеры в третичной структуре. В свернутых глобулярных белках α-спирали обычно не располагаются параллельными

Дополнение 2.3

ТОПОЛОГИЯ β-СЛОЕВ

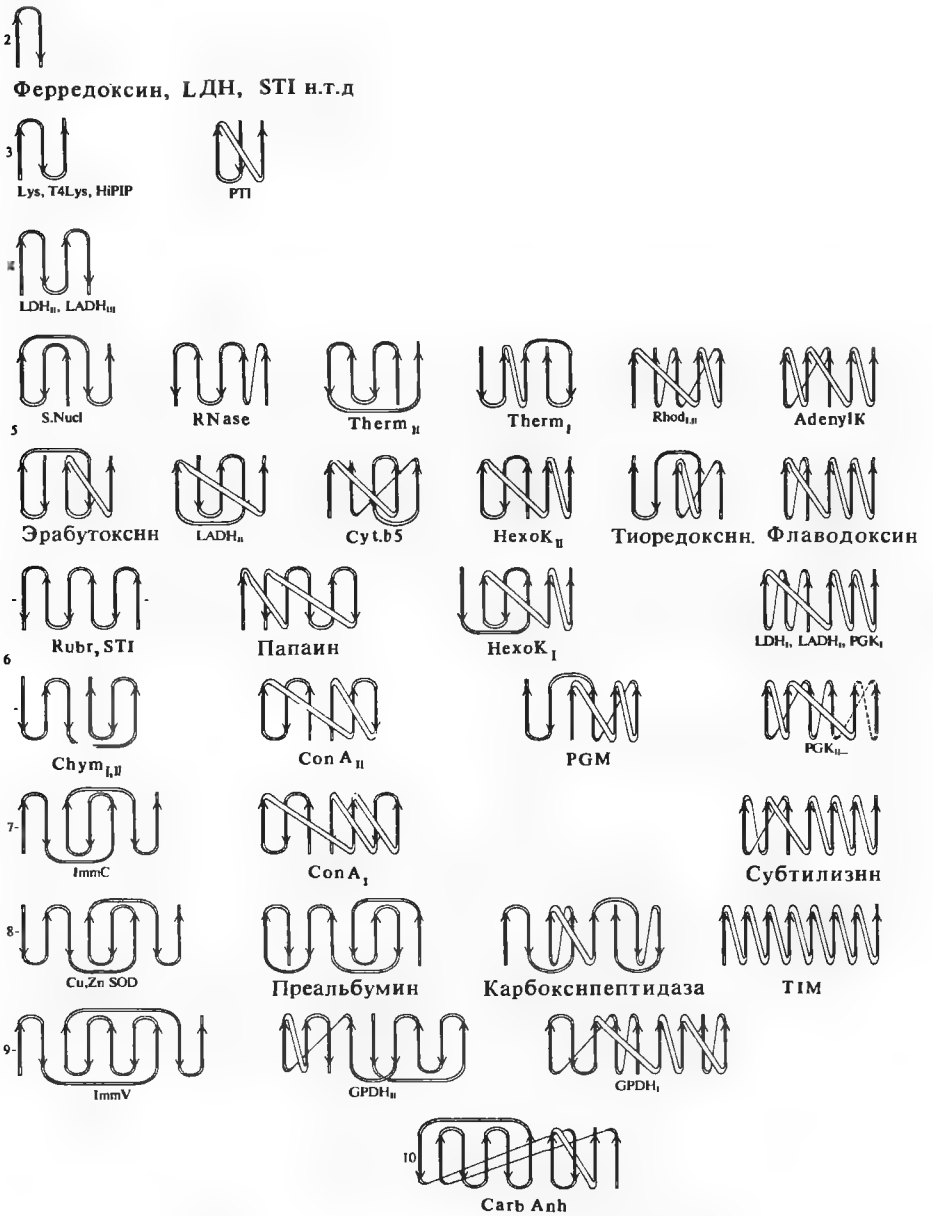
Некоторые из множества обнаруженных в белках β-слоев схематически изображены на рисунке. Каждый слой показан в проекции на плоскость, сплошными стрелками указаны направления отдельных цепей (от N- к С-концу). Соединения отдельных цепей, расположенные над плоскостью рисунка, изображены широкими двойными линиями, лежащие в плоскости — тонкими двойными линиями и, наконец, проходящие под плоскостью — тонкими одинарными линиями. Слои, напоминающие цилиндрические бочонки, обозначены справа и слева черточками. Рисунок иллюстрирует только топологию β-структур, но не дает представления о длине и конформации отдельных цепей или их соединений.

Чтобы оценить степень абстрактности изображенных топологических представлений, надо сравнить эти рисунки с более реалистическими изображениями структур карбоангидразы (см. рис. 2.34), лактатдегидрогеназы (см. рис. 2.39,*А*), лизоцима (см. рис. 2.33), супероксиддисмутазы (см. рис. 2.43) и вариабельного домена иммуноглобулина (см. рис. 2.43).

Джейн Ричардсон вывела ряд полезных обобщений из известных структур β-слоев, в том числе следующие.

1. Все перекрестные соединения (за исключением одного) являются правыми (соединения, расположенные выше плоскости β-слоя, изображены идущими в направлении от верхнего левого угла в правый нижний; соединения, расположенные ниже плоскости β-слоя, направлены в противоположную сторону). [J. S. Richardson, Proc. Natl. Acad. Sci. USA, **73**, 2619 (1976).]
2. Параллельная β-структура наблюдается только в слоях, состоящих из пяти и большего числа цепей.
3. Структуры, состоящие только из параллельных или только из антипараллельных слоев, предпочтительнее, чем смешанные структуры.
4. Топологические узлы не встречаются. Ни один из слоев, изображенных на рисунке, не завяжется в узел, если его потянуть за N- и С-концы в разные стороны.
5. N-концевой участок цепи с большей вероятностью, чем С-концевой, располагается в середине слоя. Это согласуется с представлением о том, что свертывание цепи может начинаться до окончания синтеза пептидной цепи.

Схематические структуры, изображающие β-слои, расположены в порядке возрастания числа цепей сверху вниз, а в порядке убывания отношения доли антипараллельной структуры к доле параллельной — слева направо. На рисунке используются следующие сокращения: LDH — лактатдегидрогеназа, STI — ингибитор трипсина из соевых бобов, Lys — лизоцим яичного белка, T4Lys — лизоцим фага T4, HiPIP — высокопотенциальный железосодержащий белок, PTI — ингибитор трипсина из поджелудочной железы, LADH — алкогольдегидрогеназа из печени, Therm — термолизин, Rhod — роданаза, Adenyl K — аденилаткиназа, cyt b_5 — цитохром b_5, Hexo K — гексокиназа, Rubr — рубредоксин, PGK — фосфоглицериаткиназа, Chym — химотрипсин, Con A — конканавалин А, PGM — фосфоглюкомутаза, ImmC — константный домен иммуноглобулина, ImmV — вариабельный домен иммуноглобулина, Cu,ZbSOD — Cu, Zn-супероксиддисмутаза, TIM — триозофосфатизомераза, CPDH — глицеральдегидрофосфатдегидрогеназа, CarbAnh — карбоангидроза, S. Nucl — нуклеаза из *S. aurens*, RNase — рибонуклеаза поджелудочной железы. [J. Richardson, Nature, **268**, 495 (1977).]


2
Ферредоксин, LДН, STI н.т.д
3
Lys, T4Lys, HiPIP
PTI
4
LDH$_{II}$, LADH$_{III}$
5
S.Nucl
RNase
Therm$_{II}$
Therm$_{I}$
Rhod$_{I,II}$
AdenylK
Эрабутоксин
LADH$_{II}$
Cyt.b5
НехоК$_{II}$
Тиоредоксин.
Флаводоксин
6
Rubr, STI
Папаин
НехоК$_{I}$
LDH$_{I}$, LADH$_{II}$, PGK$_{I}$
Chym$_{I,II}$
Con A$_{II}$
PGM
PGK$_{II}$
7
ImmC
Con A$_{I}$
Субтилизин
8
Cu,Zn SOD
Преальбумин
Карбоксипептидаза
TIM
9
ImmV
GPDH$_{II}$
GPDH$_{I}$
10
Carb Anh

пакетами. Не существует и определенных способов упаковки элементов соседних вторичных структур, за исключением, может быть, очень специальных последовательностей аминокислотных остатков.

Основной вопрос, возникающий при изучении третичной структуры, состоит в том, как ее рассматривать — как единое целое или как состоящую из отдельных доменов.Можно привести примеры обоих вариантов. В большинстве рассмотренных выше структур пептидная цепь пересекает структуру во всех направлениях много раз. Однако в термолизине (см. рис. 2.34) структуру можно разбить на две совершенно независимые области. Важно выяснить, могут ли эти отдельные области существовать в виде стабильных свернутых структур в отсутствие остальной части белка. Если это реально, то можно исследовать процесс их реассоциации и, следовательно, непосредственно получить некоторые сведения о специфических взаимодействиях и силах, участвующих в стабилизации третичной структуры.

Наличие доменов может отражать ход образования третичной структуры из развернутой пептидной цепи. Быстрее всего собираются в свернутую или упорядоченную структуру участки протяженной полипептидной цепи, расположенные по соседству в ее первичной структуре, так как они с большей вероятностью могут оказаться сближенными в пространстве при произвольной конформации белка. В гл. 21 мы обсудим количественные данные о динамике процесса свертывания белков. Здесь мы хотим только отметить, что если структура белка делится на домены, то последние представляют собой важные промежуточные состояния в процессе свертывания. Возможно также, что белок претерпевает конформационные изменения, при которых внутренняя структура доменов остается неизменной, а их относительное расположение заметно изменяется. Таким образом,белок, состоящий из доменов, будет скорее иметь гибкую структуру, чем белок, в котором различные участки скрепляются между собой перекрещивающимися полипептидными цепями. (В качестве примера можно привести структуру иммуноглобулина G, изображенную на рис. 1.9 и 2.10).

ОКРУЖЕНИЕ ОТДЕЛЬНЫХ ОСТАТКОВ

Данные о третичной структуре белков дают ключ к решению вопроса о силах, которые должны участвовать в их стабилизации. Здесь мы сформулируем несколько общих правил, которые были выведены при изучении многих известных третичных структур белков. В гл. 5 приведены независимые физико-химические доказательства, объясняющие или подтверждающие эти правила. Последние нельзя считать абсолютными, и исключения из них обычно связаны с функциональными свойствами.

1. Конформации отдельных пептидных групп лежат в областях наибольшей стабильности, предсказываемых на основании исследований модельных соединений и полуэмпирических энергетических расчетов. Образование вторичной и третичной структур редко происходит за счет нарушения локальной стабильности небольших групп остатков. В тех случаях, когда это имеет место, разумеется, меняются свойства этих остатков, а иногда возрастает и их роль в катализе и других функциях. Но в целом тенденция отдельных остатков находиться в их наиболее устойчивых состояниях оказывает сильнейшее влияние на то, какими могут быть вторичная и третичная структуры.

2. Заряженные остатки должны располагаться на поверхности молекулы белка (см., например, рис. 2.31). Мы уже упоминали о том, насколько невыгодно погружение внутрь молекулы зарядов. На поверхности молекулы распределение зарядов у разных белков может быть различным. Иногда оно оказывается достаточно однородным. Области чередующихся положительных и отрицательных зарядов обладают меньшей энергией, чем кластеры подобных зарядов. Однако распределение зарядов в кластерной форме встречается довольно часто, и следует ожидать, что такие третичные структуры более подвержены разрушению при пониженной ионной силе.

Аннзотропные распределення зарядов приводят к появлению у молекулы большого днпольного момента. Весьма вероятно, что определенные распределения зарядов играют важную бнологическую роль, способствуя связыванию других больших нли малых молекул или, возможно, приводя к предпочтнтельным орнентациям в снльных электрическнх полях, напрнмер полях, существующих между обеимн сторонамн некоторых мембран. Кулоновские взанмодействня янляются самыми дальнодействующими из снл, которые участвуют в сильных взаимодействиях, преобладающих в бнологических системах, н, вероятно, можно найти примеры того, как эти снлы специфическн используются для достиженя определенных функцнональных целей. В ряде случаев заряды обнаружнваются внутрн белковой молекулы, при этом они нензменно оказываются на остатках, непосредственно участвующих в процессе катализа нлн выполняющнх другие функции.

3. Гндрофобные неполярные остатки в основном должны располагаться внутрн третичной структуры глобулярного водорастворимого белка. Надо отметнть, что это правило нельзя сформулировать без некоторых оговорок. Пример, в котором оно безусловно выполняется, представлен на рис. 2.32. Но даже здесь локализация остатков не так сильно ограничена, как н случае заряженных групп. Потерн энергни при экспонированин нескольких неполярных остатков не так уж велики. В олнгомерных белках контакты между субъедницами часто образованы неполярными остатками. Однако отдельная субъеднница может быть достаточно стабильной сама по себе и иметь частично гидрофобную поверхность, необходимую для взаимодействия с другими субъединицами.

Исключения из этого правила могут иметь место в случае белков, снльно взаимодействующих с липидами. Хотя мы до снх пор не располагаем данными ни об одной третичной структуре мембранного белка нли лнпопротеида, полученными при достаточно высоком разрешенин, позволяющем увидеть отдельные боковые группы, тем не менее было бы весьма удивительно, если бы эти молекулы не имели значительного числа экспонированных наружу гндрофобных остатков. Известно несколько белков, которые совершенно нерастворимы в воде, но хорошо растворимы н функционально активны в органических растворителях. Их структуру можно было бы сравнить со структурой обычного белка, вывернутой наизнанку. Заряженные группы, по-видимому, находятся внутрн, взаимодействуя друг с другом и нспользуя для стабилнзации в случае необходимостн связанные ионы. Гидрофобные остатки концентрируются на поверхности и участвуют во взанмодействнях с растворителем.

4. Фактически все доноры и акцепторы водородных связей находятся там, где они могут образовывать эти связн. Онн лнбо располагаются на поверхности, образуя водородные связи с водой, либо встраиваются в области упорядоченной вторичной структуры, лнбо локализуются там, где они могут образовывать внутренние водородные связи в третичной структуре (это хорошо видно на рнс. 2.33). Если бы не образонывалось достаточно большое чнсло водородных связей, это могло бы оказаться слишком невыгодным энергетически. Поэтому полярные боковые группы стремятся расположиться на поверхности, так как там существует бóльшая вероятность найти партнера для образовання водородной связи. Внутренние полярные боковые цепи часто формнруют водородные связи друг с другом илн же с блнжайшнмн участкамн полнпептидного скелета.

Заметим, однако, что большинство третичных структур белков нзвестно только из результатов рентгеноструктурного анализа. Прн этом разрешенне, с которым получены рентгеновские данные, редко бывает достаточно высоким, чтобы можно было увидеть водородные связн. При интерпретацнн этих данных мы склонны преувеличивать число водородных связей, поскольку, исходя нз современных представленнй, мы думаем, что такие связи *должны быть*.

Некоторые детали организации третичных структур могут стать более яснымн, еслн изучить распределение в ннх типов боковых групп, встречающихся в α-спнралях. Еслн спнраль располагается вблнзи поверхностн, то нз самой ее структуры ясно, что боковые цепи будут направлены попеременно внутрь и наружу белка, если двигаться вдоль цепн.

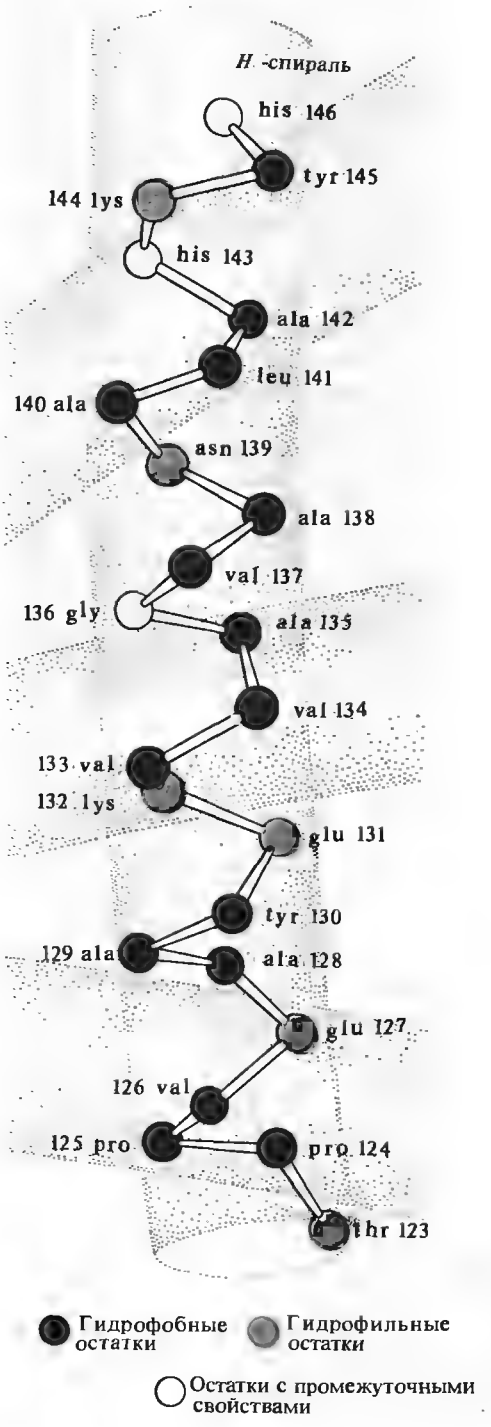


РИС. 2.36. Расположение *H*-спирали в третичной структуре β-субъединицы гемоглобина человека. Гидрофобные боковые цепи в основном локализуются вблизи поверхности *H*-спирали, там, где она контактирует с *A*-, *E*-, и *F*-спиралями, расположенными на рисунке перед ней.

Таким образом, можно ожидать, что распределение полярных и неполярных боковых групп для такой спирали будет периодическим, что и было обнаружено в действительности. В качестве примера на рис. 2.36 изображена одна из спиралей гемоглобина. Заметим, что на N-конце цепи находятся остатки пролина. Такое расположение пролинов часто встречается в белках. Может создаться впечатление, что их роль заключается в том, чтобы не позволить такой стабильной вторичной структуре, как α-спираль, продолжаться неограниченно. Полипептидная цепь в β-слое может поворачивать назад сама с помощью β-изгибов, и поэтому такая структура менее склонна выходить за пределы глобулы. Боковые цепи β-слоев направлены вверх и вниз от плоскости слоя. Если слой расположен вблизи поверхности, то распределение боковых аминокислотных цепей обычно тоже можно предсказать.

ПЛОТНОСТЬ УПАКОВКИ АМИНОКИСЛОТНЫХ ОСТАТКОВ В СВЕРНУТЫХ МОЛЕКУЛАХ БЕЛКА

Насколько плотно упакованы остатки в белке? Действительно ли он является почти сплошной структурой или же в нем есть пустые или заполненные растворителем области? В третичной структуре белковых молекул иногда находили несколько больших полостей. В миоглобине есть область, хорошо соответствующая размерам атома ксенона; это объясняет тот факт, что белок связывает один эквивалент этого инертного газа. Рибонуклеаза S имеет несколько полостей, которые, по-видимому, пусты. В лизоциме и α-химотрипсине имеется еще больше полостей, вероятно, заполненных растворителем. В кристаллической структуре α-химотрипсина было найдено 16 внутренних молекул воды [1]. Однако визуально большинство белков выглядит довольно плотными структурами, как это видно на рис. 2.29 и 2.30.

Для изучения этого вопроса можно провести количественный анализ. Один из подходов заключается в использовании точных координат атомов для вычисления плотности упаковки. Если считать, что каждый атом представляет собой сферу, радиус которой равен вандерваальсову радиусу атома, то плотность упаковки выражается отношением суммы объемов атомов в какой-либо области к полному объему, занимаемому данной областью. Это отношение для наиболее плотного расположения сфер составляет 0,74, для бесконечных цилиндров — 0,81 и для сплошного твердого тела — 1,00. Определения плотности упаковки вблизи поверхности недостаточно точны, и результаты, полученные для небольшого числа рассматриваемых атомов, не вполне одинаковы. Однако если белок разделить на средние по величине объемы и рассчитать плотность упаковки, как это показано на рис. 2.37, то обнаруживается ряд интересных особенностей.

Значение средней плотности упаковки во внутренних областях рибонуклеазы S примерно равно 0,75, что лежит в пределах, найденных для кристаллов типичных небольших органических молекул. Например, для кристаллов Gly—Phe—Gly плотность упаковки равна 0,749. Однако она немного колеблется, даже если усреднение проводится по достаточно большим областям белка. Надо отметить, что в областях, соседних с бороздкой активного центра, плотность упаковки невелика, тогда как в областях, расположенных рядом (т.е. дальше от активного центра), она намного превышает среднюю величину. Это может играть определенную функциональную роль. Более рыхлые области обладают большей гибкостью, что допускает внутренние движения, повороты боковых групп и т.д. Области с высокой плотностью упаковки, по всей вероятности, обладают очень большой

[1] Не следует путать полости и внутренний растворитель, рассматриваемые здесь, с полостью и внутренним растворителем *между* белковыми молекулами в кристалле. Такой растворитель в большом количестве находится между молекулами белков в кристаллах, однако он не рассматривается как часть третичной структуры.

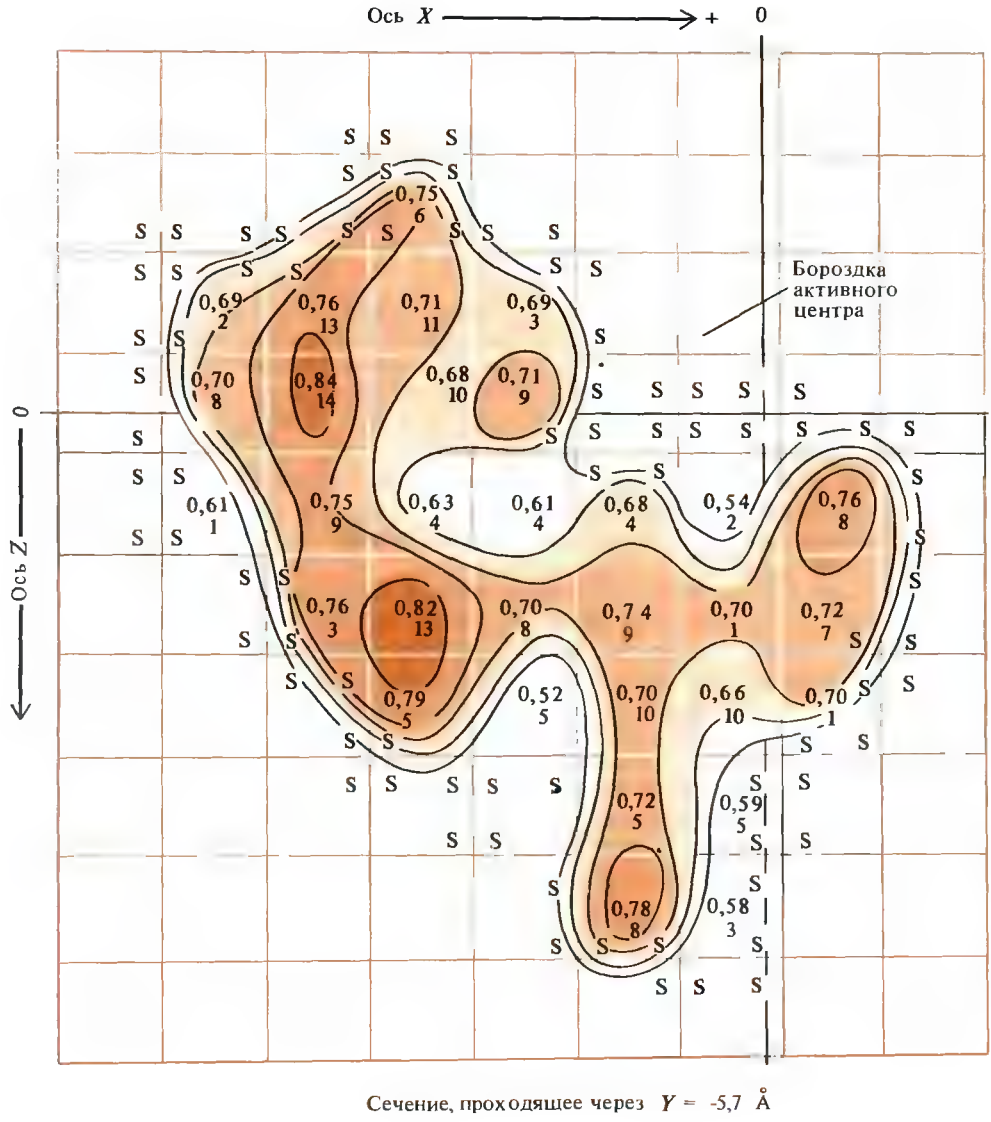


РИС. 2.37. Сечение, проведенное через центр рибонуклеазы S из поджелудочной железы быка (фрагмент содержит остатки 21 — 124). Структура помещена в кубическую решетку с длиной ребра ячейки 5,6 Å. В каждой клетке указаны количество атомов в соответствующем кубе и средняя плотность упаковки этих атомов (т.е. доля объема куба, которую они занимают). Цветом показаны примерные зоны равной плотности. Чтобы очертить поверхность молекулы, указан гипотетический слой молекул растворителя, обозначенных буквой S. [F. M. Richards, J.Mol.Biol., 82, 1(1974.)]

жесткостью. Таким образом, распределение плотности упаковки позволяет установить, какие типы движения возможны для структуры в целом. Оно указывает на некоторую подвижность остатков в области активного центра, а также на то, что в значительных перемещениях должны участвовать жесткие участки структуры. Перемещаясь как целое, эти области могут передавать конформационные изменения на значительные расстояния.

ОБЪЕМ И ПЛОТНОСТЬ БЕЛКОВ

Суммируя вандерваальсовы объемы всех атомов, можно рассчитать молекулярный объем белка. Получаемый таким методом объем равен 16 900 $Å^3$ для рибонуклеазы S (если пренебречь остатками, недостаточно хорошо идентифицированными с помощью рентгеноструктурного анализа) и приблизительно 18 100 $Å^3$ для лизоцима. Эти величины на 10% больше объемов, измеренных в растворах и полученных из данных о молекулярных массах и экспериментально определенных плотностей (см. гл. 10).

Экспериментально найденные плотности ρ (г/$см^3$) дают основание предполагать, что связанная вода имеет ту же плотность, что и остальная. Такое допущение было бы вполне приемлемо, но оно противоречит рентгеновским данным. Поверхность белковой молекулы несет много зарядов, а заряженные группы, как известно, тесно связываются с водой, увеличивая ее плотность. Этот электрострикционный эффект составляет 16 $Å^3$ на заряженную группу. Рибонуклеаза S несет 24 заряда, что может уменьшить кажущийся объем в растворе на 400 $Å^3$ на молекулу белка. Этот эффект мог бы объяснить уменьшение кажущегося объема белка, но, как мы видим, он недостаточно велик, чтобы полностью компенсировать разницу в объемах.

Проблема еще усложняется, если плотность ρ рассчитывается по координатам, найденным из рентгеновских данных, как функция расстояния от центра масс белка. Результаты, полученные для нескольких белков, приведены на рис. 2.38, где дается ρ для сферических оболочек с радиусами, увеличивающимися на 2 Å. На очень малых расстояниях величины плотности сильно флуктуируют, поскольку объемы сферических слоев (и соответственно число атомов, которые они содержат) малы. Затем функция выходит на плато, где ρ почти не зависит от расстояния. Для различных малых белков это соответствует удалению на 8-16 Å от центра масс. При большем радиусе плотность постепенно падает, что связано с неправильной формой поверхности каждого белка. В этих расчетах молекулы растворителя не учитываются.

В простых случаях, моделируя белки однородными твердыми телами, мы должны получить постоянную плотность, соответствующую действительной плотности белка в растворе. Однако она меньше на 10-25%. Это различие оказывается даже больше рассмотренного выше. Существуют три возможных объяснения. Не только влияние электрострикции, но и другие эффекты могут увеличивать плотность молекул воды, связанных на поверхности белков. Если считать, что белки окружены водной оболочкой толщиной в одну молекулу, то она должна была бы иметь плотность от 1,35 до 1,60 г/$см^3$, чтобы устранить различия между расчетной и наблюдаемой плотностью. Эти значения кажутся слишком большими, особенно если учесть, что плотность остальной воды приблизительно равна 1,0 г/$см^3$.

Второе возможное объяснение состоит в том, что внешние области белка могут иметь бóльшую плотность, чем внутренние. Из рис. 2.37 такого заключения сделать нельзя, но надо учитывать, что гидрофобные остатки, имеющие в среднем меньшую, чем у полярных остатков, плотность, стремятся расположиться внутри молекулы белка. Наличие плато при ρ от 1,1 до 1,2 г/$см^3$ легко можно было бы объяснить тем, что имеется внутренняя область, состоящая из одних лишь углеводородоподобных соединений. В действительности при этих значениях ρ число пептидных цепей может быть даже несколько бóльшим, однако для полного согласия с экспериментально полученными для растворов зна-

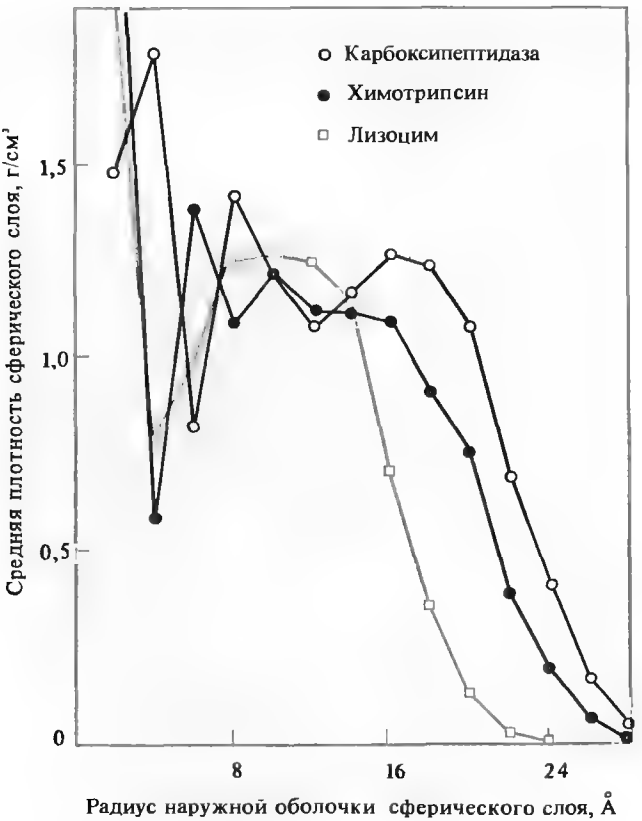


РИС. 2.38. Зависимость средней плотности сферического слоя от радиуса этого слоя для нескольких белков. Центры сферических слоев находятся в центре масс соответствующего белка. При расчете плотности, выраженной в г/см^3, учитывались все атомы, кроме атомов растворителя. [W. Kauzmann, Nature, 248, 447, (1974).]

чениями ρ плотность внешних областей белка должна была бы достигать 1,50-1,55 г/см^3. Это слишком большая величина по сравнению с известными плотностями кристаллических пептидов.

Третье объяснение заключается в том, что внутри белка в действительности находится больше молекул воды, чем следует из данных рентгеноструктурного анализа. При этом для разрешения противоречий между экспериментом и теорией требуется от 100 до 300 молекул воды на одну молекулу белка, что кажется весьма маловероятным. Таким образом, мы приходим к заключению, что каждый из этих трех механизмов, вероятно, вносит лишь некоторый вклад в решение проблемы. Этот пример служит для иллюстрации тех трудностей, с которыми мы сталкиваемся, пытаясь ответить даже на относительно простые вопросы, касающиеся третичной структуры белка.

ДИНАМИЧНА ИЛИ СТАТИЧНА ТРЕТИЧНАЯ СТРУКТУРА?

Приведенные выше расчеты плотности упаковки указывают на необходимость получения информации о гибкости различных областей белка и его проницаемости для растворителя или других малых молекул. Общих ответов на эти вопросы не существует, но некоторые методы в принципе позволяют проникнуть в суть проблемы. Движения ароматических аминокислотных боковых цепей можно обнаружить спектроскопическими методами, например, по данным о поляризации флуоресценции. В большинстве случаев результаты, полученные для этих остатков, соответствуют ожидаемым скоростям вращения белковой молекулы в целом; это дает основание предполагать, что они жестко закреплены в молекуле. Однако колебательные движения с малыми угловыми амплитудами могут в таких экспериментах и не проявляться. Прикрепленные снаружи меченые молекулы обычно столь же жестко связаны, но не ясно, не препятствует ли их взаимодействие с поверхностью белка нормальным движениям.

Метод ЯМР, особенно с использованием изотопа ^{13}C, является мощным средством исследования движения определенных боковых цепей, так как атомы некоторых типов, например ε-углеродные атомы лизина, можно регистрировать независимо. Вообще говоря, результаты указывают на то, что боковые цепи на поверхности имеют значительную вращательную свободу. Намного меньшей свободой, если она вообще есть, обладают остатки внутри молекулы. Этот результат согласуется с данными рентгеноструктурного анализа. Боковые цепи на поверхности молекул в белковых кристаллах кажутся разупорядоченными, что можно объяснить их вращательным движением. Основная масса молекулы представляется весьма жесткой, но следует помнить, что имеющиеся данные рентгеноструктурного анализа редко получаются с достаточно высоким разрешением, чтобы можно было с уверенностью об этом говорить.

Оценка проницаемости белка или доступности отдельных его остатков — очень сложный вопрос. В общем он ставится следующим образом: могут ли малые молекулы проникнуть внуть белка и участвовать в химических реакциях или изменять спектроскопические свойства внутренних остатков? Для различных малых молекул иногда получаются разные результаты. Например, молекулы кислорода, по-видимому, способны достигать внутренних областей белка и тушить флуоресценцию большинства (или даже всех) остатков триптофана в глобулярном белке. Ионы иода могут достичь только некоторых из них. Вместе с тем и те и другие одинаково эффективны в качестве тушителей флуоресценции, если третичная структура белка разрушена денатурирующими агентами. Неизвестно, обусловлено ли различие во взаимодействиях с нативными структурами зарядом иодида, его большими размерами или более тонким локальным влиянием окружения на процесс тушения.

Целый ряд ковалентных химических модификаций боковых аминокислотных цепей позволяет установить, что третичная структура защищает многие остатки, обладающие обычной реакционной способностью в свободных аминокислотах или в коротких пептидах. Если данную боковую группу можно модифицировать, то, вообще говоря, мы вправе заключить, что она экспонирована в растворитель. Если реакция протекает без потери функциональной активности, можно быть уверенным в том, что она не вызывает серьезных нарушений структуры белка. Однако если остаток не способен вступать в реакцию, опыт показывает, что нельзя с уверенностью заключить, что он погружен в структуру. Локальное окружение, даже на поверхности белка, может существенно отличаться от обычного водного окружения.

Расчеты динамики белковых структур дают многообещающий способ исследования гибкости и проницаемости белковых молекул (см. пример, приведенный в Дополнении 2.4).

Дополнение 2.4

РАСЧЕТЫ МОЛЕКУЛЯРНОЙ ДИНАМИКИ БЕЛКОВ

Динамика третичных структур белков была подробно рассмотрена в недавних работах М. Карплюса. Он использовал метод молекулярной динамики для определения подвижности остатков в структуре ингибитора трипсина, выделенного из поджелудочной железы быка. В этом методе силы, действующие на остатки, определяются из предварительно рассчитанной потенциальной функции (типа описанной в гл. 5). Сила, приходящаяся на каждый атом или остаток, равна

$$\mathbf{F}_i = -\sum_j \nabla V_{ij}$$

где V_{ij} — потенциал взаимодействия между i-м и j-м атомами.

Для свободных атомов или молекул массой m_i каждая сила, согласно закону Ньютона, вызывает ускорение $\mathbf{a}_i$:

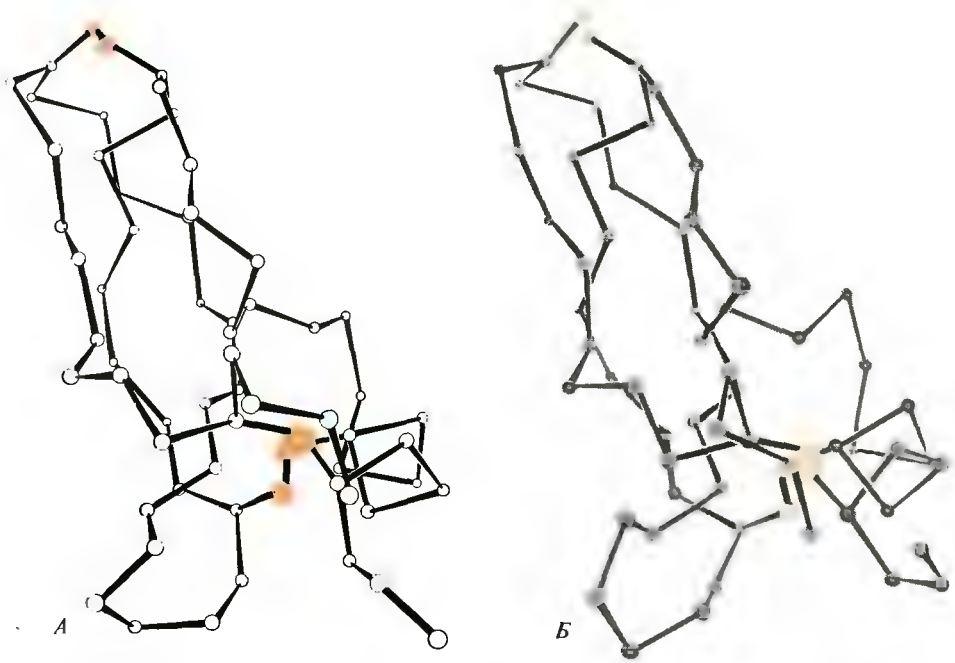
$$\mathbf{F}_i = m_i\mathbf{a}_i$$

В расчетах методом молекулярной динамики оба уравнения решают совместно для всех i атомов при малых значениях приращения времени. Затем рассчитывают новые положения атомов и повторяют все вычисления.

Результаты этих расчетов приводят к нескольким важным следствиям (см. рис.). Среднеквадратичное отклонение остатка от положения, полученного методом рентгеноструктурного анализа, составляет 1 Å. Это значительная величина, которая указывает на то, что структура как бы «дышит», делая белок проницаемым для малых молекул. Не все остатки движутся одинаково. Некоторые области обладают большей подвижностью, чем другие. Однако подвижность нескольких остатков настолько велика, что ход полипептидной цепи, найденный из рентгеновских данных, становится непоказательным.

МЕТОДЫ СРАВНЕНИЯ ТРЕТИЧНЫХ СТРУКТУР

Третичные структуры многих белков определены методом рентгеноструктурного анализа. Было бы весьма полезно, хотя это и трудно сделать, провести тщательное сравнение таких структур между собой. Если третичные структуры очень похожи, визуальное наблюдение позволяет расположить их в одинаковой ориентации; и при этом сразу можно обнаружить отдельные различия. Сравните, например, структуры нескольких сериновых эстераз или α- и β-субъединиц гемоглобина и молекулы миоглобина. Чем сильнее различия, тем труднее обнаружить сходство структур. Становится невозможно найти ориентации двух молекул, которые облегчили бы сравнение. Компьютер может методом проб и ошибок найти ориентации, при которых структуры имеют максимальное сходство, но это иногда приводит к потере интуитивного чувства структуры у исследователя. Для



А. Структура ингибитора трипсина из поджелудочной железы быка, полученная методом рентгеноструктурного анализа; C^{α}-атомы показаны незакрашенными кружками, атомы серы в остатках цистина — закрашенными кружками. *Б.* Структура молекулы рассчитана для молекулярных движений, происходящих за $3 \cdot 10^{-12}$ с.

сравнения необходимо определенное абстрактное или просто обобщенное представление структур. Имеются по меньшей мере два подхода к этой проблеме, оказавшиеся полезными. В обоих особое внимание уделяется элементам вторичной структуры, поскольку их относительно легко идентифицировать визуально.

Наиболее часто используется метод, состоящий в построении простейшей схемы, отражающей основные особенности вторичной структуры молекулы и взаимосвязь составляющих ее элементов. Эти схемы не похожи на реальные структуры, однако их легко описать и запомнить. На рис. 2.39 представлен набор таких схем для белков с центрами, связывающими нуклеотиды. Структурные сходства и различия схем совершенно очевидны. Они могут служить основой для последующего более детального и более строгого сравнения структур с использованием пространственных моделей или координат атомов.

Другой метод для сравнения вторичных структур более косвенный, но он не требует нахождения ориентации, в которой структуры взаимно совместимы. Основная идея этого метода заключается в использовании для сравнения не координат атомов, а расстояний между ними. Например, можно вычислить все расстояния между C^{α}-атомами и разделить их на две (или большее число) категории, характеризующие степень близости остат-

ков в третичной структуре. При этом теряется большая часть информации о структуре, поскольку межатомные расстояния определяют ее не так полно и точно, как координаты самих атомов. Однако получающиеся сводные данные легко представить в виде двумерной карты (рис. 2.40). По вертикальной и горизонтальной осям на рисунке отложены номера аминокислотных остатков в полипептидной цепи, начиная с N-конца. Точки, характеризующие расстояния между парами остатков, локализованных недалеко друг от друга в первичной последовательности, оказываются расположенными вблизи диагонали карты. Естественно, пары таких остатков чаще всего будут близко расположены и в третичной структуре.

Для каждого белка строится своя карта. Сходные особенности вторичной структуры четко выделяются на такой карте. Например, α-спирали выглядят прямыми с тангенсом угла наклона, равным -1, смещенными относительно диагонали на три остатка, так как водородные связи между первым и четвертым остатками в α-спирали требуют очень близкого расположения C^{α}-атомов первого и третьего остатков. Все β-слои также выглядят прямыми линиями. Тангенс угла наклона прямой, соответствующей параллельным β-

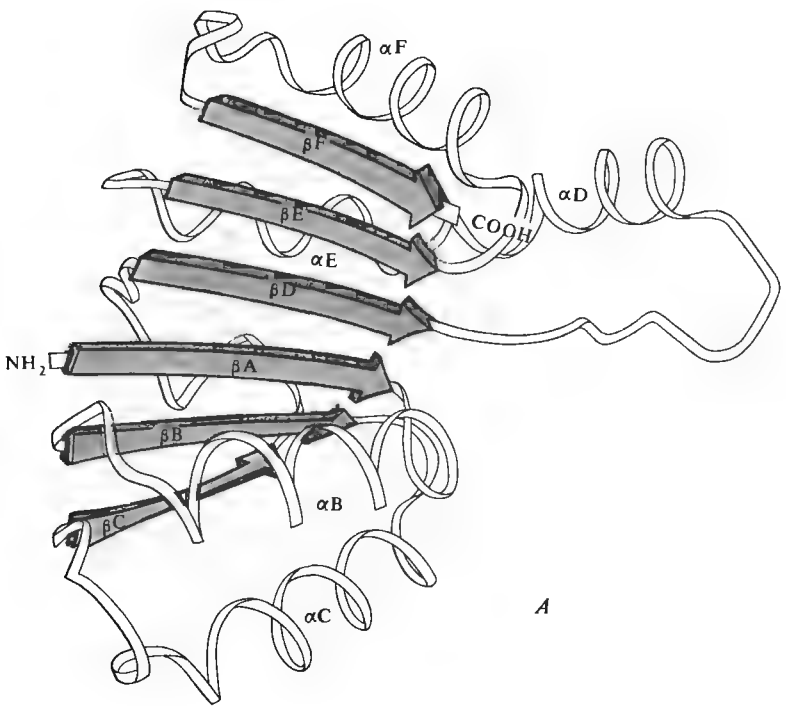
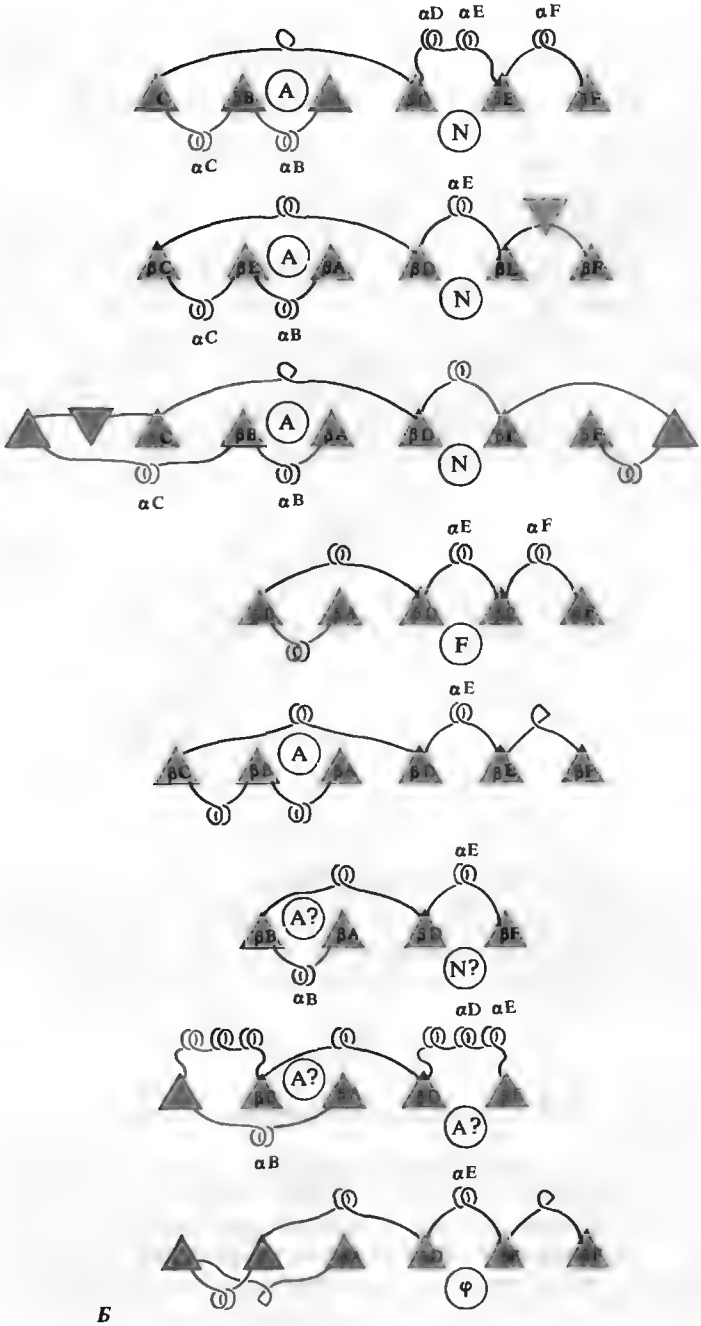


РИС. 2.39. Нуклеотидсвязывающие домены белков и некоторые родственные им третичные структуры. *А*. Схематическое изображение части третичной структуры лактатдегидрогеназы. Стрелками указано направление цепей в β-слое от N-конца к С-концу. *Б*. Схематические изображения сходных моделей третичных структур лактатдегидрогеназы и других белков. Треугольниками показаны β-слои, расположенные таким образом, что их N-концы находятся ближе к наблюдателю. Витки обозначают α-спирали. Буквы A,N и F указывают положения известных мест связывания аденина в молекулах NAD или ATP, никотинамида в NAD и флавина соответственно. [M. G. Rossmann In: The Enzymes, 3d ed., vol. 9, ed. P. Boyer (New York: Academic Press, 1975), p.61.]


αD αE αF
A
N
αC αB
Лактатдегидрогеназа
Растворимая
малатдегидрогеназа
αE
Алкогольдегидрогеназа
из печени
Глицеральдегид –
фосфатдегидрогеназа
αE αF
F
Флаводоксин
Фосфоглицераткиназа
A?
N?
Роданаза
αD αE
Аденилаткиназа
φ
Субтилизин

Б

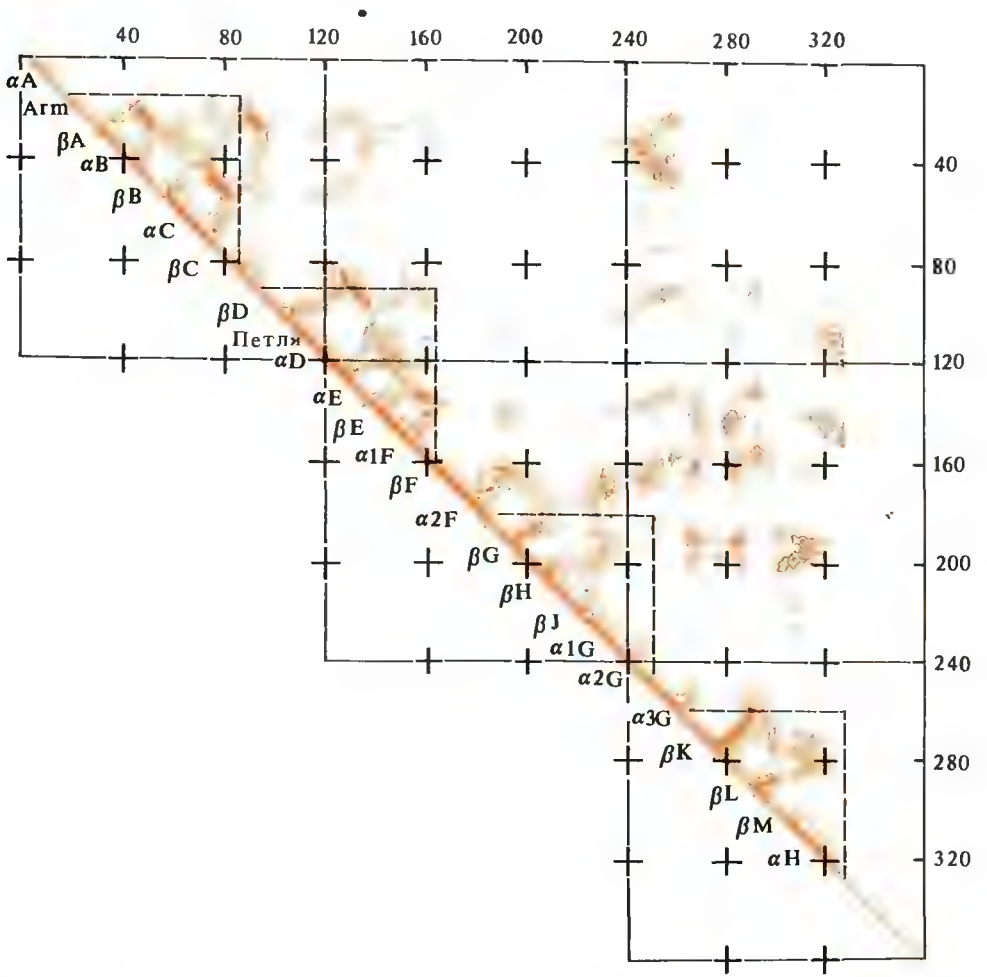


РИС. 2.40. Диаграмма расстояний между парами C^{α}-атомов в лактатдегидрогеназе. Изолинии проведены через пары атомов, находящихся на расстояниях 0; 4; 8; 12; и 16 Å друг от друга. По горизонтальной и вертикальной осям отложены номера остатков. На рисунке изображена только половина матрицы контактов, поскольку вторая половина тождественна первой ($r_{ij} = r_{ji}$). Четыре структурных домена обведены пунктирными линиями. [A. Liljas, M. G. Rossmann, Ann. Rev. Biochem., **43**, 475, (1974).]

слоям, равен -1, но она смещена дальше от диагонали. Для антипараллельных слоев тангенс угла наклона равен $+1$.

Схематические структурные карты могут помочь выявить сходство в структурных особенностях и облегчить сравнение действительных координат. Они также весьма полезны для выяснения того, свертываются ли цепи с образованием доменов. Домены выступают на карте как группы (кластеры) близких контактов, далеко отстоящие от других таких же групп. Из характера карты каждого домена можно судить о том, похожи ли их внутренние структуры. Например в лактатдегидрогеназе можно видеть четыре домена

(см. рис. 2.40). Первые два имеют сходную структуру, и, как выяснилось, они сходны и в функциональном отношении. Этот фермент связывает динуклеотид NAD. Каждый из этих двух доменов служит для связывания одной мононуклеотидной части кофермента. Схематические структуры этих доменов приведены на рис. 2.39.

СВЯЗЬ МЕЖДУ СТРУКТУРНЫМ И ФУНКЦИОНАЛЬНЫМ ПОДОБИЕМ

В качестве простейшего случая структурного подобия можно взять две (или большее число) молекулы белка с довольно сходными структурами и близкими функциями. Однако особый интерес представляет сравнение таких белков, которые явно сходны либо только в структурном, либо только в функциональном отношении. Сравнение целого ряда белков, связывающих нуклеотиды, дает замечательный пример этого. Два упомянутых выше домена, связывающие NAD в лактатдегидрогеназе, были обнаружены также в растворимой малатдегидрогеназе, в глицеральдегид-3-фосфатдегидрогеназе и в алькогольдегидрогеназе печени. На рис. 2.39 приведено схематическое изображение вторичной структуры этих областей. В перечисленных выше четырех белках не отмечается какой-либо значительной гомологии в первичных структурах. Более того, положения динуклеотид-связывающих доменов в первичной структуре совершенно различны. В трех белках этот домен располагается в области первых 150 остатков с N-конца, тогда как в алкогольдеги-

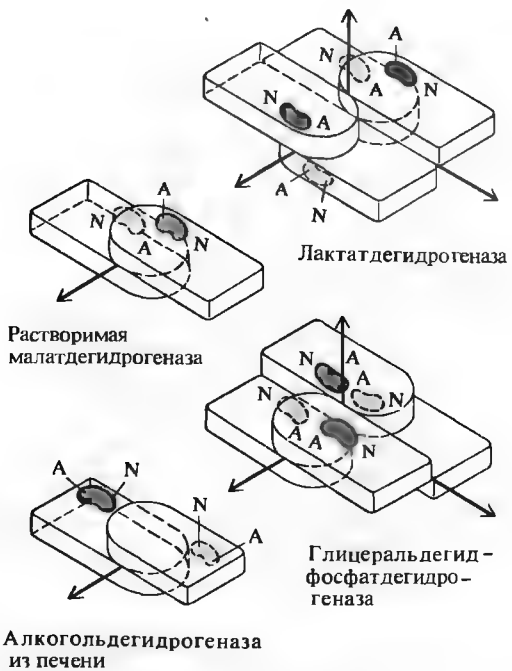


РИС. 2.41. Схематическое изображение четвертичных структур четырех белков, содержащих сходные области связывания NAD. Оси симметрии показаны стрелками. (M.G.Rossmann et al. In: The Enzymes, 3d ed., vol. 9, ed. P. Boyer (New York: Academic Press, 1975), p.61.]

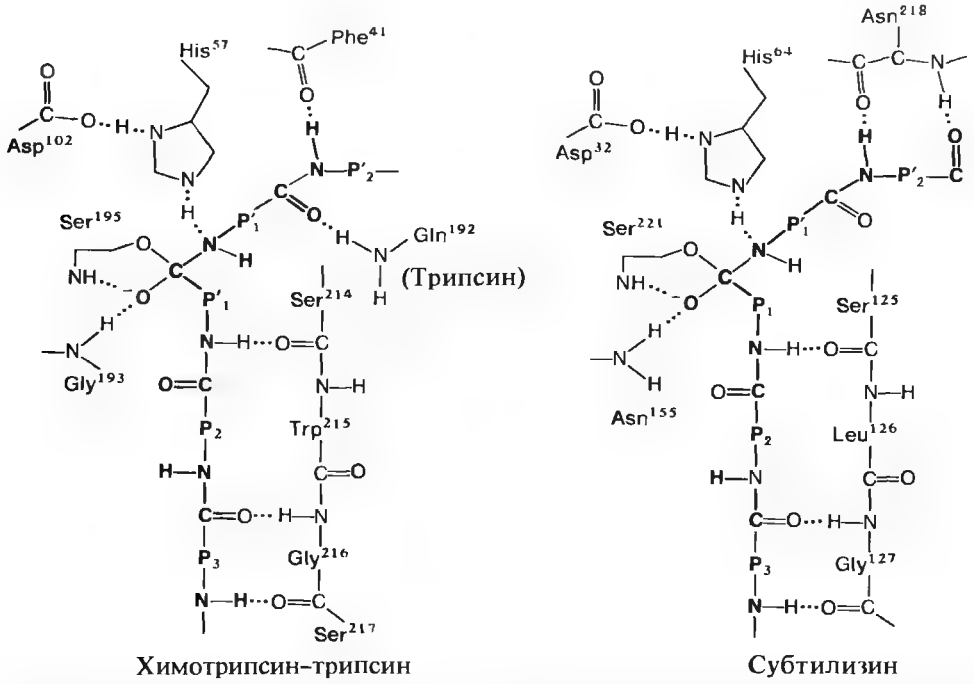


РИС. 2.42. Сходство областей связывания субстрата в химотрипсине и субтилизине. Пептидная цепь субстрата показана жирным шрифтом. Обратите внимание, что помимо остатков Asp^{102} и Asp^{32}, Ser^{195} и Ser^{221}, His^{57} и His^{64}, участвующих в каталитическом процессе, сходным образом в обоих белках расположены и другие остатки, способствующие связыванию пептидной цепи субстрата. [F.Liljas, M.G.Rossmann, Ann.Rev.Biochem., 43, 475 (1974).]

дрогеназе печени он начинается со 196-го остатка. Положения доменов в четвертичной структуре белков также не согласуются между собой. Как показано на рис. 2.41, в глицеральдегид-3-фосфатдегидрогеназе четыре области связывания NAD образуют в четвертичной структуре один кластер, тогда как в лактатдегидрогеназе они разделены на две пары.

Весьма примечательно, что такие же или очень сходные структурные домены обнаружены в белках, не связывающих NAD. Например, флаводоксин, связывающий FMN, содержит один домен, эквивалентный одному из двух доменов лактатдегидрогеназы. Фосфоглицераткиназа имеет два таких домена, и один из них участвует в связывании ADP. Аденилаткиназа, по-видимому, тоже содержит сходные домены, так же как и роданаза. Даже субтилизин, протеолитический фермент, насколько известно, не взаимодействующий с нуклеотидами, имеет в своей третичной структуре домен типа, обнаруженного в лактатдегидрогеназе.

Видимо, природа, найдя эффективный способ построения области связывания, возвращалась к нему снова и снова. При этом возникает ряд интересных вопросов, на которые пока нет ответа. Обусловлено ли это сходство только эволюцией? Имели ли все такие белки в далеком прошлом одного общего предка? Если это действительно было так, то последующая дивергенция структуры и функций стерла все сходство, кроме подобия третичной структуры этих доменов. Может быть, сходство доменов обусловлено простыми структурными ограничениями, налагаемыми на белки? Возможно, что существует не так

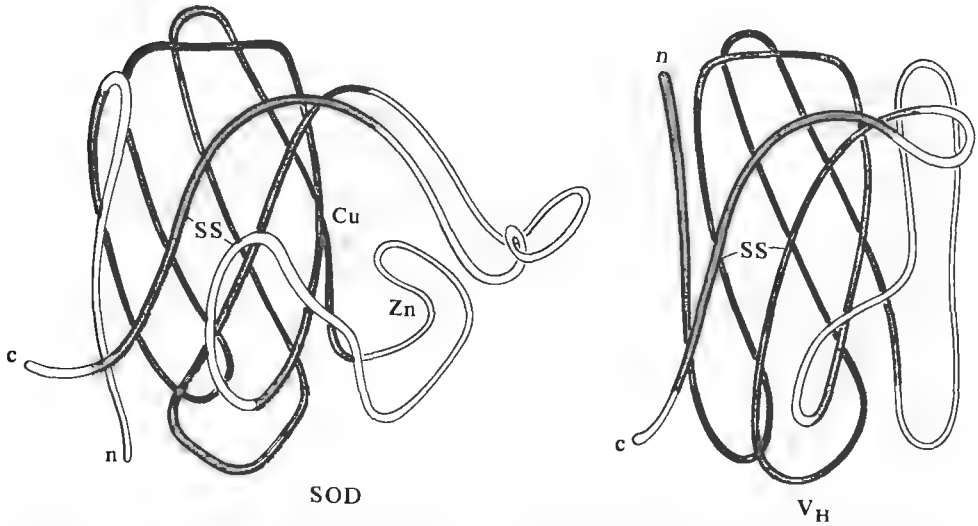


РИС. 2.43. Сравнение схематического представления структур бычьей супероксиддисмутазы (SOD) и вариабельного домена тяжелой цепи иммуноглобулина G(V_H). Основная часть структуры — β-бочонок (показан в цвете) — состоит из семицепочечного β-слоя и в обоих белках почти одинакова. (Рисунок предоставлен Д. Ричардсоном.)

уж много способов свертывания пептидной цепи, приводящих к образованию термодинамически стабильной конформации области связывания. Быть может, такие домены обладают какими-то уникальными физическими особенностями — структурой, гибкостью, жесткостью и т.д., что делает их особенно подходящими или эффективными для подобной функциональной роли? Ответить на поставленные выше вопросы можно будет только после детального изучения физических свойств этих доменов и сил, которые их стабилизируют.

Семейство сходных или эквивалентных нуклеотид-связывающих доменов было обнаружено М. Россманом в результате целенаправленного поиска близких структур, которые могли бы обусловить возникновение родственных или почти тождественных функциональных центров. Заметим, что этот домен — только часть третичной структуры каждого белка. В другом интересном случае речь идет об участии лишь очень небольшой доли всей белковой молекулы. Каталитические центры химотрипсинового семейства протеаз чрезвычайно похожи на каталитический центр субтилизина (рис. 2.42). Их сходство существует несмотря на то, что общий характер свертывания полипептидных цепей в обоих белках совершенно различен. Эквивалентность этих участков предполагает, очевидно, подобие каталитических механизмов. Здесь возникают вопросы, связанные с ролью дивергенции и конвергенции в эволюции белков. Может оказаться, что существует очень небольшое число каталитических механизмов, доступных белкам, которые не используют для функционирования кофакторы или простетические группы. Это может быть обусловлено тем, что 20 обычных аминокислот реализуют лишь ничтожную долю известных химических возможностей, заложенных в малых органических молекулах.

Вероятно, наиболее удивительный из известных случаев структурного подобия — сходство третичной структуры супероксиддисмутазы и одного домена Fab-фрагмента иммуноглобулина G. Как следует из рис. 2.43, общий вид третичной структу-

ры обеих молекул (за исключением одной области) практически тождествен. В настоящее время мы не можем предложить никакого объяснения такому сходству. Эти два белка выполняют совершенно разные функции. В самом деле, участки цепи, ответственные за каталитическую активность супероксиддисмутазы, и связывающие антиген центры IgG структурно различны. Не избрала ли природа какой-то основной стабильный структурный каркас, который она сохраняет неизменным даже в молекулах с совершенно разными первичными структурами и лишь незначительно модифицирует для достижения совершенно различных функциональных целей? Или это сходство отражает то обстоятельство, что существует сравнительно мало способов, которыми сегменты пептидных цепей могут свертываться в пространстве, образуя компактные глобулярные структуры? И, наконец, последняя возможность, которая ставит трудную задачу перед биохимиками: не являются ли одии или оба этих белка полифункциональными и не обладают ли они иными видами активности (в дополнение к известным), достаточно близкими, чтобы обусловить сходство их структур?

2.6. Четвертичная структура

Многие белки, выделенные из клетки, состоят из нескольких субъединиц, нековалентно связанных между собой. По своему составу, сложности и организации субъединицы могут сильно различаться. В простейшем случае белок состоит только из двух тождественных субъединиц. Эти субъединицы имеют одинаковую первичную и третичную структуру, а в рамках четвертичной структуры связаны между собой осью вращательной симметрии. В качестве примера можно привести алкогольдегидрогеназу из печени, третичная структура которой уже обсуждалась. Белки, состоящие из различных субъединиц, организованы несколько сложнее. Например, укажем на гемоглобин ($\alpha_2\beta_2$) и РНК-полимеразу из *E.coli* ($\alpha_2\beta\beta'\sigma$). Известиы белки с еще более сложной субъединичной организацией. Некоторые имеют большое число субъединиц различных типов, общий размер которых и стехиометрия строго фиксированы, например пируватдегидрогеназа $[t_{24}(p_2)_{12}(f_2)_{12}]$. Другие белки представляют собой полимерные структуры, состав которых может быть постоянным, а общие размеры меняются, например микротрубочки $[(\alpha\beta)_n]$. При исследовании таких белков необходимо ответить на следующие вопросы. Из каких субъединиц они построены? Каково пространственное расположение субъединиц? Что стабилизирует такую организацию субъединиц? Как субъединицы собираются in vivo? Каковы функциональные причины наличия в белках субъединиц? Ответы на некоторые из этих вопросов мы дадим, а остальные обсудим в дальнейшем.

СООБРАЖЕНИЯ СИММЕТРИИ

Для четвертичных структур с небольшим числом субъединиц можно сделать некоторые предположения относительно типов структуры на основе соображений о симметрии молекул. В табл. 2.8 приведены данные о субъединичном составе для нескольких типичных белков. Как видно из этой таблицы, число тождественных субъединиц часто оказывается кратным 2 или 3, а не бóльшим простым числам. Это следствие наиболее вероятных симметричных способов упаковки твердых тел в трехмерном пространстве. (Обзор операций симметрии приведен в Дополнении 2.5.)

Рассмотрим простейший пример белков, состоящих из субъединиц одного типа, организованных в одну из идеальных групп точечной симметрии. Приводимые ниже примеры всегда можно обобщить иа более сложные случаи. Каждая субъединица, вообще говоря, — асимметричное образование. Более того, все составляющие ее α-углеродные атомы также асимметричны. Поэтому четвертичная структура белков не может содер-

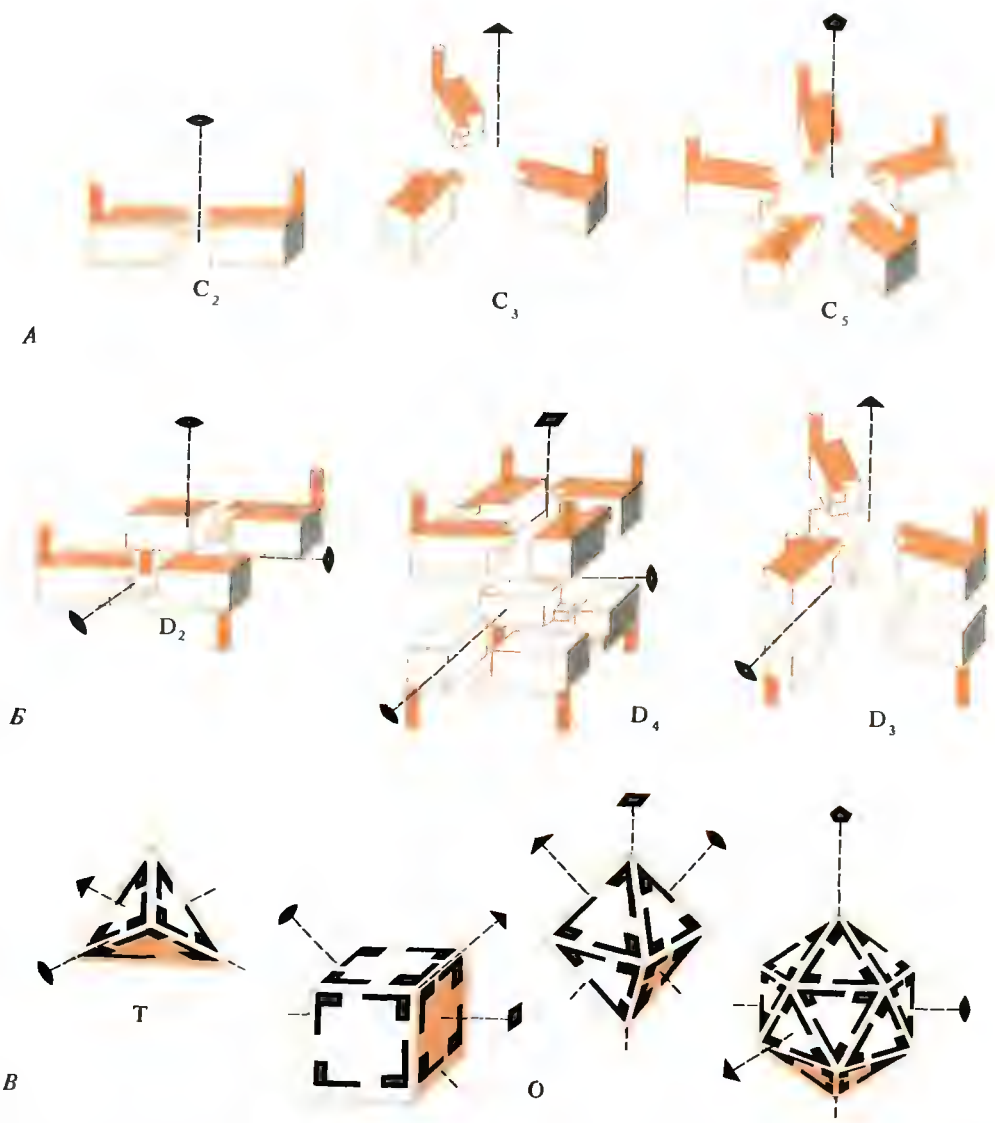


РИС. 2.44. Возможные группы точечной симметрии для наборов одинаковых субъединиц. *А*. Ансамбли субъединиц с аксиальной симметрией. Показано положение осей вращения для структур с симметрией C_2, C_3 и C_5. *Б*. Ансамбли с диэдрической симметрией. Показано положение взаимно перпендикулярных осей симметрии для структур с симметрией D_2, D_3 и D_4. *В*. Ансамбли с кубической симметрией. Показаны некоторые оси симметрии в кубе и в структурах с тетраэдрической (*Т*), октаэдрической (О) и икосаэдрической симметрией. (Рисунки Ирвинга Гейса.)

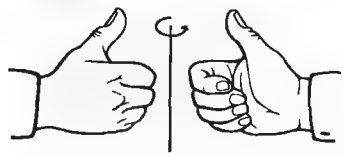
жать ни плоскостей зеркального отражения, ни инверсионных осей. Единственно возможными для белков операциями точечной симметрии являются вращения. Все возможные способы расположения субъединиц, обладающие только осями вращения, распадаются на три класса групп точечной симметрии (рис. 2.44).

1. *Аксиальная симметрия.* Эти структуры обладают только одной осью вращения. Пусть n — число субъединиц; тогда поворот на $360/n$ градусов переводит структуру саму в себя. Следовательно, все центры субъединиц можно расположить в вершинах правильного многоугольника. Группы точечной аксиальной симметрии обозначаются либо n, либо C_n. Фактически в олигомерных белках проявляется симметрия C_2 и C_3.

Дополнение 2.5

ОПЕРАЦИИ СИММЕТРИИ

Операциями симметрии являются повороты, параллельные переносы или другие преобразования, при которых объект не измеияется. Операции точечной симметрии представляют собой особый класс преобразований, которые не только оставляют объект неизмеиным, но и оставляют по крайней мере одну точку в объекте неподвижной. Например, поворот на 180° меняет местами две тождественные субъединицы, и его можно представить следующим образом:

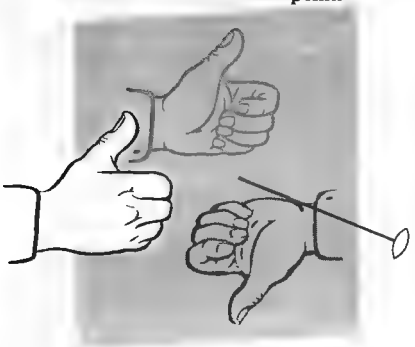


Ось симметрии C_2

Такому повороту соответствует ось симметрии C_2. Это пример операции точечной симметрии, поскольку точка, расположенная на оси, остается при этом преобразовании неподвижной. Симметричное расположение двух субъединиц можно получить с помощью операции зеркального отражения. Это также пример операции точечной симметрии, поскольку часть объекта, расположенная на самой плоскости, остается в результате отражения неподвижной.

Зеркальное отражение

Операция инверсии соответствует одновременному повороту вокруг оси симметрии C_2 и отражению в плоскости, перпендикулярной этой оси. Это также является операцией точечной симметрии. Неподвижной остается при этом единственная точка, находящаяся на пересечении оси вращения и плоскости отражения. Эта точка называется точкой инверсии.

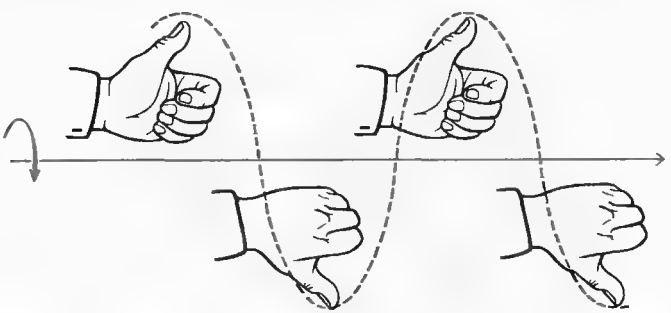


Инверсия

Группе точечной симметрии данного объекта соответствует совокупность всех операций симметрии, которые оставляют объект неизменным. Полностью асимметричному объекту соответствует группа симметрии, называемая группой E. К такой группе симметрии относится только тождественное преобразование, которое оставляет объект неизменным и неподвижным. Другие примеры операций точечной симметрии приводятся в тексте. К операциям симметрии относятся параллельные переносы, а также возможные вращения и некоторые другие операции. Они называются операциями пространственной симметрии. В качестве примера рассмотрим простой параллельный перенос вдоль оси.

Параллельный перенос

В результате комбинации параллельного переноса с вращением получается винтовая ось. Она образует спираль или «ленту». На рисунке изображено сочетание вращения вокруг оси симметрии C_2 с параллельным переносом вдоль этой оси. Дополнительный материал по вопросам, связанным с симметрией, можно найти в гл. 13. Операции симметрии прекрасно изложены в книге П. Эткинса. [P. W. Atkins, Physical Chemistry, San Francisco, W. H. Freeman and Company, 1978. Имеется перевод: Эткинс П. Физическая химия. В 2-х т. — М.: Мир, 1980.]



Винтовая ось

Таблица 2.8

СУБЪЕДИНИЧНЫЙ СОСТАВ БЕЛКОВ, СОДЕРЖАЩИХ БОЛЕЕ ОДНОЙ СУБЪЕДИНИЦЫ[1)]

Белок	Источник	Молекулярная масса белка	Число субъединиц	Молекулярная масса субъединиц
Фактор роста нервов	Мышь	26 518	2	13 259
Лютеинизирующий гормон	Овца	27 322	1	12 500
			1	14 830
Ингибитор химотрипсина	Картофель	39 000	4	9800
Супероксиддисмутаза	*E.coli*	39 500	2	21 600
Гемоэритрин	*Phascolosoma*	40 600	3	12 700
Галактокиназа	Человек	53 000	2	27 000
Гемоглобин	Млекопитающее	64 500	2	16 000
			2	16 000
Tu·Ts-комплекс	*E.coli*	65 000	1	41 500
			1	28 500
Малатдегидрогеназа	Крыса	66 300	2	37 500
Авидин	Куриное яйцо	68 300	4	18 000
Тропонин	Кролик	80 000	1	37 000
			1	24 000
			1	20 000
Щелочная фосфатаза	*E.coli.*	86 000	2	43 000
Прокарбоксипептидаза А	Крупный рогатый скот	88 000	1	40 000
			2	23 000
Серил-тРНК-синтетаза	*E.coli*	100 000	2	50 000
Нуклеозиддифосфокиназа	Дрожжи	102 000	6	17 000
Тубулин	Поросенок	110 000	1	56 000
			1	53 000
Лактатдегидрогеназа	" "	140 000	4	35 000
Триптофансинтетаза	*E.coli*	148 000	2	45 000
			2	28 700

2. *Диэдрическая симметрия.* Этот тип симметрии встречается при наличии хотя бы одной оси вращения C_2, перпендикулярной другой оси вращения n-го порядка. Минимальное число субъединиц в олигомере равно $2n$, где n-любое натуральное число. По принятой номенклатуре, эти группы точечной симметрии обозначаются D_n, где n-наибольший порядок оси вращения, имеющейся в группе. Другая иоменклатура предлагает точное обозначение всех осей симметрии. Например, группе D_2 соответствует обозначение 222, поскольку иаличие двух взаимно перпеидикулярных осей второго порядка автоматически порождает третью, перпендикулярную им ось второго порядка (см. рис. 2.44). Группе D_3 присвоено обозначение 32, группе D_4 — 422.

3. *Кубическая симметрия.* Правильный куб имеет три оси симметрии четвертого порядка (проходящие через центры противоположных граней), четыре оси третьего порядка (вдоль диагоналей, связывающих противоположные вершины) и шесть осей второго порядка (проведенных через центры противоположных ребер). Точечные группы вращения, относящиеся к кубической симметрии, имеют по крайней мере четыре оси третьего порядка. К ним относятся тетраэдрическая, октаэдрическая и икосаэдрическая группы. На рис. 2.44 приведены некоторые оси симметрии, которые могут иметь место в подобных структурах.

Таблица 2.8

Белок	Источник	Молекулярная масса белка	Число субъединиц	Молекулярная масса субъединиц
lac-Репрессор	То же	160 000	4	40 000
Метионин-тРНК-синтетаза	" "	170 000	2	85 000
Репликаза фага $Q\beta$	" "	205 000	1	70 000
			1	65 000
			1	45 000
			1	35 000
Ариламидаза	Человек	223 500	6	38 100
Лейцинаминопептидаза	Свинья	255 000	4	63 500
Изоцитратдегидрогеназа	Дрожжи	300 000	8	39 000
Аспартаттранскарбамоилаза	*E.coli*	310 000	6	33 000
			6	17 000
Нитрогеназа	*Clostridium*	330 000	2	59 500
			4	27 500
			2	50 700
Енолаза	*T.aquaticus*	355 000	8	44 000
Глутаминсинтетаза	*Neurospora*	360 000	4	90 000
РНК-полимераза (кор)	*E. coli*	400 000	2	39 000
			1	155 000
			1	165 000
Апоферритин	Лошадь	443 000	24	18 500
Глутаминсинтетаза	*E.coli*	592 000	12	48 500
Изоцитратдегидрогеназа	Крупный рогатый скот	670 000	16	41 000
Овомакроглобулии	Куриное яйцо	650 000	2	325 000
Гемоглобин	*Arenicola*	2 850 000	48	54 000
Пируватдегидрогеназный комплекс	*E.coli*	5 000 000	24	91 000
			24	65 000
			24	56 000

[1]D. W. Darnall, I. M. Klotz, Arch. Biochem. Biophys., 166,651 (1975).

Простое тетраэдрическое расположение четырех тождественных субъединиц не обладает законченной тетраэдрической (T) симметрией. Это связано с тем, что отдельная субъединица в вершине тетраэдра асимметрична и поэтому оси симметрии C_3 отсутствуют. Таким образом, тетраэдр, в вершинах которого расположены субъединицы, имеет только симметрию D_2. Для построения системы субъединиц, имеющей симметрию T, необходимо по крайней мере 12 тождествеиных мономеров, образующих треугольники с локальной симметрией C_3 в каждой вершине тетраэдра. Аналогично, октаэдрическое расположение отдельных субъединиц обладает только симметрией D_3. Для построения законченной октаэдрической (O) группы симметрии в каждой вершине должны располагаться четыре субъединицы, образующие квадрат. Для реализации этой симметрии требуется 24 субъединицы. Заметим, что октаэдр имеет четыре оси третьего порядка, три оси четвертого порядка и шесть осей второго порядка, т.е. тот же набор элементов симметрии, что и истинный куб. Для создания системы субъединиц с кубической симметрией также необходимо 24 субъединицы; образуя треугольники, они располагаются по три субъединицы в каж-

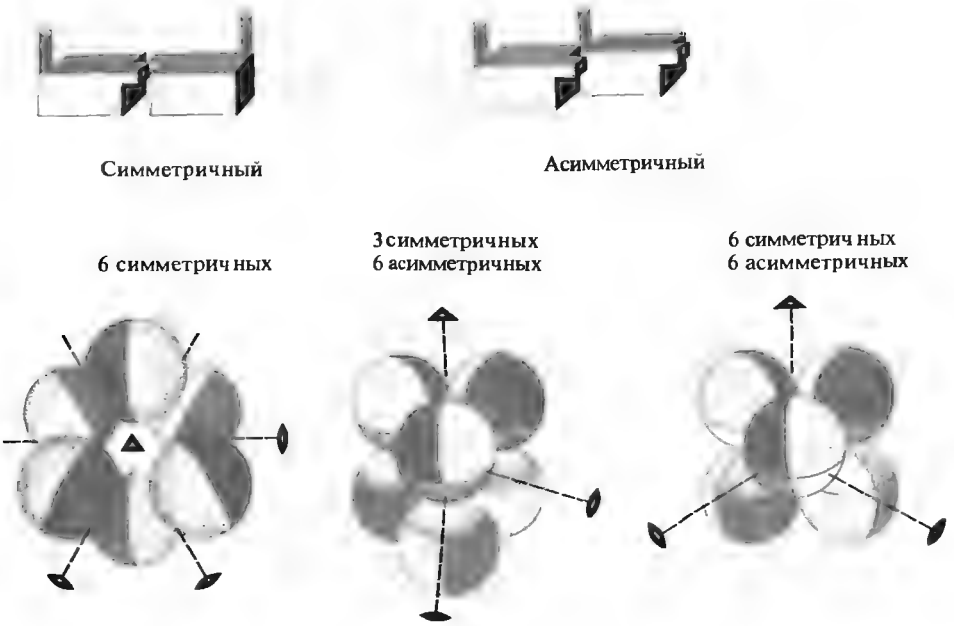


РИС. 2.45. Симметрия контактов между субъединицами. *А*. Схематическое изображение симметричного и асимметричного контактов. *Б*. Три структуры с симметрией D_3, причем каждая содержит различные наборы симметричных и асимметричных контактов. Плотность упаковки субъединиц возрастает, полное число контактов увеличивается от 6 до 12. (Рисунок Ирвинга Гейса.)

дой вершине куба. Сходным образом можно показать, что законченная икосаэдрическая симметрия потребует наличия 60 тождественных субъединиц.

Существование столь большого числа тождественных субъединиц, необходимых для построения групп кубической симметрии, практически маловероятно. Несколько возможных примеров приведено в табл. 2.9. Однако большинство олигомерных белков, обладающих хоть какой-то симметрией, попадают в классы либо с аксиальной, либо с диэдрической симметрией. Структурный смысл групп точечной симметрии становится более понятным при рассмотрении реальных контактов между поверхностями субъединиц. Существуют два возможных типа контактов между субъединицами: симметричный и асимметричный (рис. 2.45). Первый предполагает наличие двух тождественных контактных областей в субъединицах, связанных осью симметрии C_2. Асимметричный контакт предполагает, что контактные области в субъединицах различны и, следовательно, эти области для каждой субъединицы остаются доступными для образования олигомеров более высокого порядка. Таким образом, при асимметричных контактах, естественно, образуются ленты и спирали. Если шаг спирали равен нулю и геометрия контакта правильна, последовательные асимметричные контакты создают олигомеры с аксиальной симметрией, порядок которой больше двух.

Иногда различным типам субъединичных контактов может соответствовать одна и та же группа точечной симметрии. На рис. 2.45 представлены три группы из шести субъединиц, причем каждая из последних обладает симметрией D_3. Однако одна структура имеет

Таблица 2.9

ПРИРОДА КОНТАКТОВ МЕЖДУ СУБЪЕДИНИЦАМИ[1)]

Белок	Симметрия	Контактирующие области	Число вандерваальсовых контактов	Число водородных связей	Число пар ионов
α-Химотрипсии	C_2	A	443	9	1
		B	57	6	–
Конканавалин А	D_2	II к A	–	2	–
		II к B	142	14	6
		II к C	174	14	–
Гемоглобин					
Оксигемоглобин	C_2	$\alpha_1\beta_1$	110	5	–
		$\alpha_1\beta_2$	80	1	–
		$\alpha_1\alpha_2$	–	–	–
		$\beta_1\beta_2$	–	–	–
		$\alpha_1\beta_1$	98	5	–
Дезоксигемоглобин	C_2	$\alpha_1\beta_2$	69	1	1
		$\alpha_1\alpha_2$	–	–	2
		$\beta_1\beta_2$	–	–	1
Инсулин	D_3	OP	111	8	–
		OQ	99	2	1

[1)] A. Lilias, M. Rossmann, Ann. Rev. Biochem., 43, 485 (1974).

два различных набора из трех симметричных контактов. Вторая, с более плотной упаковкой, имеет три симметричных и шесть одинаковых асимметричных контактов, образующих два треугольника. Последняя группа упакована наиболее плотно. В ней имеются 12 коитактов: шесть асимметричных и шесть симметричных — по три с двумя разными типами симметрии.

Данные о симметрии весьма важны при описании субъединичных структур. Однако при очень подробиом исследовании могут наблюдаться иебольшие, но важные в функциональном отиошении отклонения от строгой симметрии даже для номиально идентичных субъединиц. Это может проявляться в виде небольших трансляционных сдвигов одной из субъединиц в четвертичной структуре или в некоторых неэквивалентных изменениях в третичной структуре. Хотя большинство белков, состоящих из множества субъединиц, образует симметричную систему, мы не можем a priori установить, что для этого требуется. Например, гексокиназа, выделенная из дрожжей, состоит из двух тождественных субъединиц, образующих димер без точечной симметрии: два ее мономера связаны друг с другом винтовой осью, но геометрия их такова, что спиральная цепочка из многих субъединиц не образуется (рис. 2.46).

АНАЛИЗ ЧИСЛА СУБЪЕДИНИЦ И ПОЛИПЕПТИДНЫХ ЦЕПЕЙ

Наиболее достоверную информацию о расположении субъединиц в четвертичной структуре дают рентгеноструктурные и электронно-микроскопические данные. Для этого часто оказывается достаточным выполнить рентгеноструктурные исследования с низким

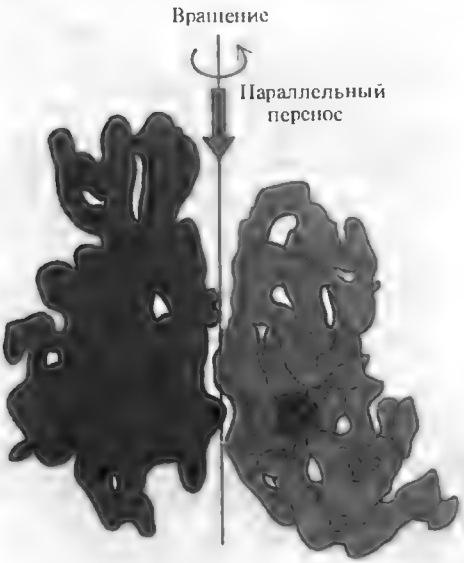


РИС. 2.46. Структура гексокиназы из дрожжей, полученная при низком разрешении. Две субъединицы связаны винтовой осью. Многогранник в правой субъединице представляет собой глюкозу. (Рисунок предоставлен Т. Стейцем.)

разрешением. Применение для отдельных типов субъединиц специфических красителей, например соответствующих антител, значительно увеличивает информативность электронно-микроскопических данных.

Электроино-микроскопические и рентгеноструктурные данные могут быть дополнены или даже заменены результатами, полученными с помощью более простых физических и химических методов. Например, с помощью гидродинамических методов, описанных в гл. 11 и 12, можно определить молекуляриую массу интактиой четвертичной структуры (M_q) с точностью 1—10% в зависимости от метода определения и присутствия неаминокислотных составных частей[1].

Во многих случаях можно предполагать, что каждая субъединица образована одной полипептидной цепью. Если это не так, то можно подобрать условия, при которых все индивидуальные полипептидные цепи в четвертичной структуре разделяются. Для разделения обычно используют сильные денатурирующие агенты, такие как додецилсульфат иатрия, и обработку восстанавливающими агентами для разрушения всех дисульфидных связей. Часто все получающиеся остатки цистеина алкилируют иодуксусной кислотой или иодацетамидом для предотвращения восстановления дисульфидных связей при последующих измерениях. Методы разделения почти всегда дают возможиость определить число различных типов полипептидных цепей. Например, с помощью электрофореза часто разделяют белки, которые абсолютио идентичны, за исключением единичных зарядовых различий. Молекулярную массу каждой из денатурированных цепей (M_c) можно определить гидродинамическими методами, хотя и не всегда с такой же точностью, как в случае нативных белков. В качестве альтернативы используют химические методы для определения числа аминокислот, приходящегося на один N-конец. Вместе с аминокислотиым анализом это дает возможность достаточно точно определить молекулярную массу цепи.

Непосредствеиный переход от молекулярной массы полипептидной цепи к молекуляриой массе субъединицы не всегда прост. Если существуют цепи только одного типа,

[1] В процессе измерений надо быть уверенным в том, что не происходит значительной диссоциации на субъединицы, поскольку при диссоциации величина ошибки может возрасти на порядок.

тогда — исключая те случаи, когда субъединицы состоят из нескольких цепей, сшитых между собой дисульфидными связями, — можно установить прямое соответствие между одиой цепью и одной субъединицей. Число субъединиц тогда равно просто отношению M_q/M_c. Для белков, состоящих из полипептидных цепей нескольких типов, необходимо знать не только молекулярную массу каждой субъединицы, но и долю (W_c) по массе, приходящуюся на каждый тип цепи в четвертичной структуре. Тогда число цепей каждого типа будет равно $W_c M_q/M_c$. Этот расчет похож на расчет при простом количествением анализе, но ииогда именно здесь лежит осиовиая трудность. Например, одна или несколько субъединиц могут быть нестабильными или может оказаться, что их трудно удержать в растворeином состоянии, и тогда такие субъединицы легче, чем другие, упустить из виду в процессе разделения или анализа.

Если субъедииицы состоят из полипептидных цепей нескольких типов, то после их разделения возиикает вопрос: какую из цепей приписать той или ииой субъединице? Обычно оказывается иеобходимым вернуться к иативиой структуре и попытаться, используя более мягкие условия деиатурации, разделить ее на субъединицы, сохраняющие нативную третичную структуру. Затем после очистки субъединиц определяют молекулярную массу и аминокислотный состав субъединиц каждого типа. Например, при добавлении сильного денатурирующего агента гемоглобин распадается на две α- и две β-цепочки. В условиях мягкой денатурации образуются главным образом $\alpha\beta$-димеры; свободные α- и β-субъединицы встречаются очень редко. Основываясь на аминокислотном составе цепей, гемоглобин можно было бы иазвать четырехсубъединичным белком, но в реакциях он ведет себя как белок, состоящий из двух субъединиц, каждая из которых представляет собой двухцепочечный $\alpha\beta$-димер. Весьма удобный химический метод анализа четвертичной структуры — сшивание субъединиц. В случае простых олигомеров, у которых все субъединицы тождественны, белок обрабатывают раствором с избытком молекул диметилсуберимидата:

Эта молекула сшивает остатки лизина, сохраняя при нейтральном pH положительные заряды на месте положительио заряженной аминогруппы. В результате получается

Многие из этнх сшивок образуются в пределах одной субъединицы и остаются «невидимыми» для большинства аналитических методов. При образовании же сшивок между субъединицами формируются комплексы с молекулярной массой, превышающей массу одной субъединицы вдвое, втрое и т.д.; такие комплексы легко выявляются при обработке белка сильными денатурирующими агентами. Наибольшее полученное значение дает нижний предел числа субъединиц при условии, что раствор достаточно разбавлен, чтобы можно было пренебречь межмолекулярным взаимодействием.

В ряде случаев белок находится в динамическом равновесии с диссоциированными субъединицами. Это осложняет применение описанных выше методов, но облегчает использование одного дополнительного метода, который иногда оказывается весьма эффективным. Допустим, что существуют два варианта гомоолигомерного белка. Это могут быть белки разных видов, белки из различных тканей [например, лактатдегидрогеназа из сердца (Н) и мышц (М) цыпленка] или нативный и химически модифицированный препараты. Здесь очень важно, чтобы эти два варианта белка несли разные заряды. Пусть для каждого вида белка будет наблюдаться, например, следующее равновесие: $M_4 \rightleftarrows 4M$ и $H_4 \rightleftarrows 4H$. Смешивание двух видов молекул приводит к образованию целого набора ги-

бридиых структур. В тех случаях, когда самосборка каких-либо определенных гомоолигомеров не является предпочтительной, распределеиие гибридов будет биномиальным. Если начальные количества M_4 и H_4 были одинаковы, то в равновесии должно наблюдаться следующее соотношение между продуктами: $H_4(1) : H_3M(4) : H_2M_2(6) : HM(4) : M_4(1)$. Если скорость установлення равновесия не слишком велика, можно благодаря различиям в зарядах мономеров M и H выделить каждый компонент смеси методом электрофореза. Число наблюдаемых в этом эксперименте полос на 1 больше числа субъединиц.

ВЗАИМНОЕ РАСПОЛОЖЕНИЕ СУБЪЕДИНИЦ

Иифорлацию о взаимном расположеиии субъединиц можно получить и с помощью методов исследования растворов белков. Для достаточно симметричных белков, состоящих из субъединиц одного или двух типов, лучше использовать слабые денатурирующие агенты для получения фрагментов, содержащих более чем одну субъединицу каждый. Гомоолигомеры с аксиальной симметрией имеют только один тип контактов между субъединицами. В интактном олигомере на субъединицу приходится один контакт, тогда как любой фрагмент молекулы будет иметь только n-1 контакт на n субъединиц. Эти два обстоятельства означают, что диссоциация такого олигомера, вызванная любым воздействием, ослабляющим контакты между субъединицами, происходит по принципу «все или ничего». Иными словами, вероятность обнаружения значительных концентраций частично диссоциированных олигомеров очень мала.

Напротив, олигомерные белки, состоящие из субъединиц нескольких типов (или имеющие контакты нескольких типов), довольио часто диссоциируют поэтапно. Исходя из данных о составе и организации таких промежуточных ассоциатов, можно иногда сделать вывод об особенностях расположения субъединиц в исходной четвертичной структуре. Например, в гемоглобине часто встречаются превращения типа $\alpha_2\beta_2 \rightleftarrows 2\alpha\beta$. В димерах связываются между собой всегда одни и те же α- и β-субъединицы, откуда следует, что один тип α-β-контакта более стабилен, чем другой. Это указывает на то, что тетрамер имеет симметрию C_2. Ряд структур, согласующихся с симметрией C_2 и составом $\alpha_2\beta_2$, представлен на рис. 2.47.

В более сложных случаях, например в случае белков с шестью идентичными субъединицами, выделение тримеров сразу же исключает возможность образования миогих типов структур. Для агрегатов, состоящих из субъединиц различиых типов, определение промежуточиых ассоциатов становится особенно важным, поскольку оно может указать, какие субъедииицы находятся в контакте друг с другом, хотя, разумеется, желательно иметь более прямые данные об их взаимиом расположении. Один из возможных подходов заключается в более точном определении формы агрегата. В иекоторых случаях для этого используют гидродииамические методы (особенно если есть указания иа то, что изолированные субъединицы имеют приблизительно ту же третичную структуру, что и в интактном агрегате). Тогда можно аиализировать только *изменение* гидродинамических свойств при сборке. В гл. 10 мы покажем, что это позволяет с достаточной точностью различать конформации гемоглобина типа линейной, плоской квадратиой и тетраэдрической. Эта процедура, однако, не столь чувствительна, чтобы можно было отличить друг от друга промежуточные состояния.

Второй подход заключается в применении сшивающих агентов. В случае сложных агрегатов анализ результатов сшивания часто бывает довольно сложным делом, поэтому удобнее использовать такие сшивки, которые при желании можно разорвать, как, например, сшивки, образуемые восстанавливающим реагеитом *бис*-метилтиобутиримидатом

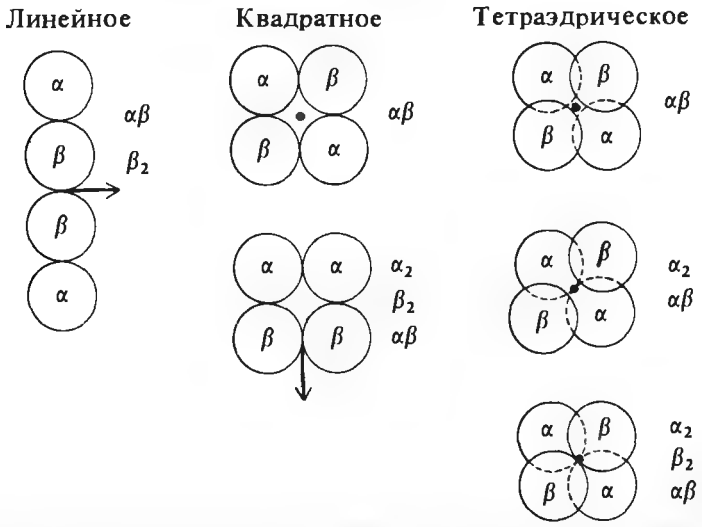


РИС. 2.47. Некоторые возможные способы расположения субъединиц в гемоглобине, совместимые с осью симметрии C_2 (направление оси указано стрелкой или цветным кружком). Другие расположения можно получить, меняя местами α и β-субъединицы. Справа от каждой структуры представлены типы возможных димеров, которых можно ожидать при сшивании гемоглобина «короткими» реагентами.

Сохранение биологической активности при сшивании двух определенных субъединиц указывает на то, что они либо соприкасаются друг с другом, либо удалены друг от друга не более чем на длину сшивающего агента. Особенно надежный контроль — выделение специфически сшитой пары субъединиц и последующая сборка ее с немодифицированными субъединицами в четвертичную структуру, в которой все молекулы содержат одну и ту же поперечную связь. Если получившиеся молекулы сохраняют биологическую активность, значит, сшивка вряд ли вызывает какие-либо серьезные искажения структуры.

На рис. 2.47 показаны примеры возможных сшивок для различных гипотетических структур гемоглобина. Очевидно, что многие структуры можно исключить, если получить хороший набор субъединиц, сшитых достаточно «короткими» реагентами. Ни один из перечисленных выше методов не может без привлечения дополнительной информации дать определенной модели расположения субъединиц даже для такого простого белка, как гемоглобин. Однако совместное использование всех трех методов позволяет исключить из рассмотрения большинство возможных структур.

Существует несколько способов, позволяющих прямо измерять расстояния между определенными точками макромолекул в растворе. Для расшифровки взаимного расположения субъединиц полезными могут оказаться два из них — метод нейтронного рассеяния и метод изучения синглет-синглетного переноса энергии. Более подробно эти методы обсуждаются в гл. 8 и 14 соответственно.

ЧЕТВЕРТИЧНАЯ СТРУКТУРА АСПАРТАТ-ТРАНСКАРБАМОИЛАЗЫ

Полезно рассмотреть один пример более подробно. Изучение аспартат-транскарбамоилазы (АТСазы) из *E.coli* весьма поучительно, поскольку эта молекула имеет наибо-

лее сложную субъединичную организацию из всех белков с известной третичной и четвертичной структурой. Фермеит катализирует следующую реакцию:

$$^{2-}OPO_3-CO-NH_2+Asp \rightleftarrows NH_2-CO-NH-\underset{\displaystyle COO^-}{\underset{|}{CH}}-CH_2-COO^- + HOPO_3^{2-} + H^+ \qquad (2.4)$$

Это первый шаг в биосиитезе пиримидинов. Интересная особенность фермента заключается в том, что присутствие пиримидина CTP зиачительно снижает скорость реакции, катализируемой АТСазой, при любой постоянной концентрации аспарагина. Это ингибирование по механизму обратной связи представляет собой биологически важиый регуляторииый процесс. Интересно, что иигибирование происходит на первом этапе биосинтеза, благодаря чему предотвращается бесполезиое (или даже вредиое) иакопление промежуточиых продуктов реакции.

Нативная АТСаза ведет себя при ультрацентрифугировании как частица с коистантой седиментации 11,3S. Комбинируя различные гидродинамические измереиия, удалось получить иезависимыми методами значения мол. массы $3 \cdot 10^5$ и $3{,}1 \cdot 10^5$. Это очень хорошее согласие между результатами[1].

Дж. Герхарт и Г. Шахман использовали *n*-гидроксимеркурибензоат для расщепления АТСазы на субъединицы. Этот реагент ковалентио связывается с сульфгидрильными группами цистеина. Данный метод обычно полезеи в тех случаях, когда стерически недоступные в интактной четвертичной структуре сульфгидрильные группы в изолированных субъединицах оказываются экспонированными в растворитель. При ультрацентрифугировании при достаточно высоких концентрациях ртути были обнаружены два типа частиц. Они имели констаиты седимеитации, равные 5,8S и 2,8S, отиошение их масс приблизительно равнялось 2 : 1. При исследовании изолированных субъединиц было сделано чрезвычайно интересное наблюдение. Субъединицы 5,8S обладали полной каталитической активностью в реакции (2.4), но не ингибировались CTP. Субъединицы 2,8S оказались лишеиы каталитической активиости, но могли связывать CTP. Поэтому субъединицы первого типа были названы каталитическими (C), а второго — регуляторными (R). При смешивании субъединиц C- и R-типов регенерировались исходные 11,3S-молекулы и восстанавливалась регуляторная функция CTP.

Были определены молекулярные массы изолированных субъединиц. Мол. масса C-субъединицы оказалась приблизительно равной $9{,}6 \cdot 10^4$, а R-субъединицы — $3 \cdot 10^4$. Затем нашли простейшую комбинацию молекулярных масс, согласующуюся с молекулярной массой нативной АТСазы и отношениями масс C- и R-субъединиц:

2 C	Общая масса	$1{,}92 \cdot 10^5$
4 R	Общая масса	$1{,}20 \cdot 10^5$
Нативная C_2R_4	Общая масса	$3{,}12 \cdot 10^5$

Согласие с наблюдаемой молекулярной массой нативной АТСазы оказалось весьма хорошим; отношение масс субъединиц C:R по крайней мере качественно было также вполне приемлемо. Для обеспечения максимальной симметрии системы было выдвинуто предположение, что каждая каталитическая субъедииица имеет два эквивалентиых места связывания с R-субъединицей. Это должно означать, что каждая 5,8S-субъединица содержит

[1] Эти величины получены из уравнения Шераги — Манделькерна и методом равновесного ультрацентрифугирования.

две одинаковые цепочки с мол. массой $4{,}8 \cdot 10^4$. Таким образом, был предположен цепочечный состав c_4r_4.

Однако на самом деле ситуация оказалась более сложной. При исследовании первичных структур было обнаружено, что мол. масса r-цепочки равна $1{,}7 \cdot 10^4$, а c-цепочки — $3{,}3 \cdot 10^4$. Это значит, что С-субъединица должна иметь состав c_3 и общую мол. массу $9{,}9 \cdot 10^4$, а R-субъединица — состав r_2 и общую мол. массу $3{,}4 \cdot 10^4$. Комбинируя эти величины для получения молекулярной массы нативного фермента, находим:

2 C	Общая масса	$1{,}98 \cdot 10^5$
3 R	Общая масса	$1{,}02 \cdot 10^5$
Нативная C_2R_3	Общая масса	$3{,}00 \cdot 10^5$

Цепочечный состав фермента при этом будет c_6r_6. Такая модель также согласуется с имеющимися гидродинамическими данными, но значительно отличается от предложенной ранее.

О справедливости этой модели свидетельствуют и два дополнительных экспериментальных факта. Было обнаружено несколько типов кристаллов АТСазы. В одной из форм фермент обладает осью кристаллографической симметрии третьего порядка, в другой — осью второго порядка, а в третьей, наиболее информативной форме молекула АТСазы имеет в кристалле $32(D_3)$-симметрию. Это значит, что число субъединиц каждого типа в четвертичной структуре должно быть кратно шести. Химический анализ указывает на существование шести ионов Zn^{2+} в интактной 11,3S-молекуле. Каждый ион Zn^{2+} связывается с одной r-цепочкой. Эта информация наряду с данными о молекулярной массе убеждает в том, что цепочечный состав нативного белка соответствует формуле c_6r_6.

Индуцированную ртутью диссоциацию АТСазы можно представить в виде $c_6r_6 \rightarrow 2c_3 + 3r_2$. В ряде изящных экспериментов с растворами Шахман с сотрудниками задались целью выяснить, сохраняют ли отдельные c_3- и r_2-субъединицы свою целостность в нативной четвертичной структуре. Некоторые из полученных ими результатов иллюстрирует рис. 2.48. Для доказательства того, что 5,8S-субъединица состоит из трех цепочек, c_3-субъединица была обработана янтарным ангидридом; при этом образовался c_3^s — вариант субъединицы с измененным зарядом. В водном растворе между c_3- и c_3^s-формами не устанавливалось равновесия, приводящего к образованию гибридов; значит, субъединичная структура стабильна и мономер не способен к обмену. Обработка c_3 и c_3^s двухмолярным раствором гуанидингидрохлорида — сильного денатурирующего агента — приводила к образованию c- и c^s-мономеров с коэффициентом седиментации 1,9S. Смешивание этих мономеров и последующее удаление денатурирующего агента приводило к образованию четырех продуктов реакции. Они были идентифицированы как c_3, c_2c^s, cc_2^s и c_3^s (см. рис. 2.48,*А*).

Для доказательства сохранения целостности двух c_3-элементов в нативной структуре АТСазы c_3- и c_3^s-субъединицы смешивались с регуляторными субъединицами для восстановления нативной четвертичной структуры. Если бы в процессе сборки структура c_3 нарушалась, то должно было бы наблюдаться семь продуктов: c_6r_6, $c_5c^sr_6$, $c_4c_2^sr_6$ и т.д. Методом электрофореза их, вероятно, разделить нельзя. Во всяком случае, используя значения биномиальных коэффициентов, легко показать, что чистые c_6r_6 и $c_6^sr_6$ будут составлять только 1/64 часть общего количества продуктов. В действительности было обнаружено три разных продукта (в отношении 1:2:1), которые идентифицировали как c_6r_6, $c_3c_3^s$ и $c_3^sr_6$. Отсюда следует, что отдельные c_3-элементы структуры сохраняют свою целостность в четвертичной структуре белка (рис. 2.48,*Б*).

Аналогичные эксперименты были выполнены на r-субъединицах. Анализ оказался более сложным по двум причинам. Обработанные ртутью r_7-субъединицы не являются ста-

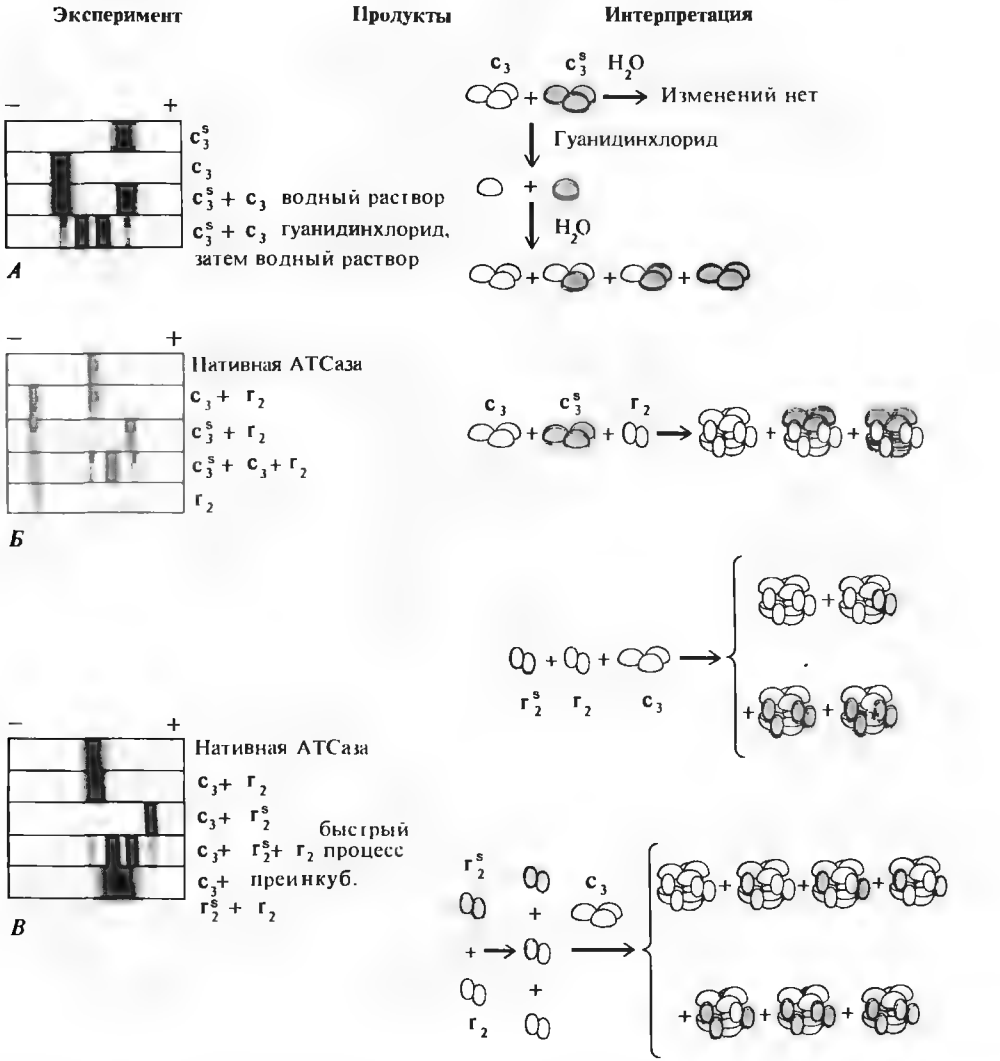


РИС. 2.48. Анализ расположения субъединиц аспартат-транскарбамоилазы (АТСазы), проведенный с помощью электрофореза смесей нативных и обработанных янтарным ангидридом субъединиц. На каждом рисунке, схематически изображающем реакцию, обработанные янтарным ангидридом субъединицы показаны в цвете. *А*. Субъединицы c_3 и c_3^s сохраняют свою целостность в водном растворе, однако диссоциируют при добавлении гуанидинхлорида. c_3-субъединица содержит три с-цепи. *Б*. c_3 и c_3^s сохраняют свою целостность в процессе реконструкции нативной АТСазы, происходящей с добавлением r_2. В нативной АТСазе содержатся два элемента c_3. *В*. r_2 и r_2^s сохраняют свою целостность в процессе реконструкции, происходящей с добавлением c_3, но обмениваются субъединицами при инкубированин в отсутствие c_3. В нативной АТСазе содержатся три r_2-субъединицы. [H.C. Schachman. In: The Harvey Lectures, vol. 68 (New York: Academic Press, 1974).]

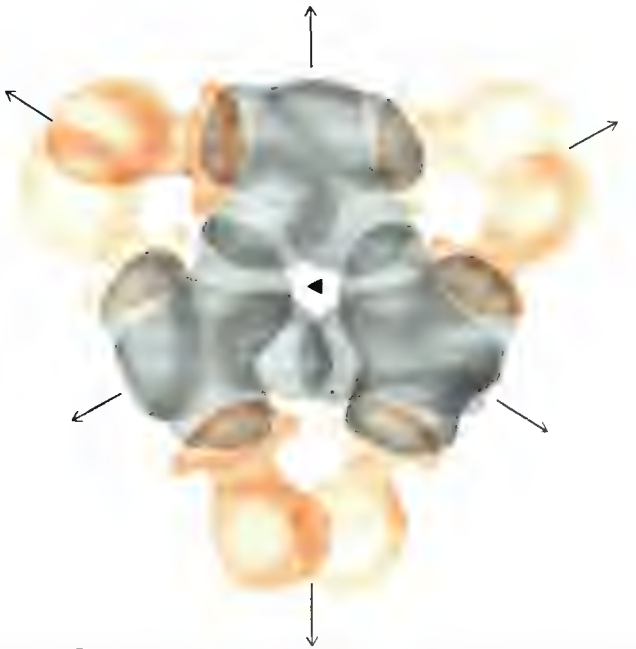


РИС.2.49. Структура АТСазы, определенная методом рентгеноструктурного анализа. Каталитические субъединицы расположены выше и ниже регуляторных (цветных) субъединиц. Ось симметрии C_3 показана черным треугольником, стрелки указывают на положение осей симметрии C_2. Более подробную информацию можно найти в литературе. [H.L. Monaco, J.L. Crowford, W.N. Lipscomb, Proc. Natl. Acad. Sci. USA, **75**, 5276 (1978).]

бильными димерами, а существуют в динамическом равновесии с мономерами. Поэтому первоначально полученная мол. масса для r_2 была равна только $3 \cdot 10^4$. Для исходной формы, содержащей ион Zn^{2+}, можно найти более стабильные условия, и действительная мол. масса, определенная методом ультрацентрифугирования, оказалась равной $3{,}38 \cdot 10^4$. Обработка отдельной r_2-субъединицы янтарным ангидридом приводила к образованию r_2^s, однако при добавлении c_3 реассоциация с образованием нативной структуры АТСазы не происходила. По-видимому, остаток лизина, весьма важный для сборки, отсутствовал.

Чтобы обойти эту трудность, нативный c_6r_6-элемент был обработан янтарным ангидридом в надежде на то, что таким образом будет защищен важный остаток лизина. Метод оказался удачным, и выделенные таким образом r_2^s-субъединицы оказались способными к реассоциации. Однако в условиях, в которых проходила сборка, все еще наблюдалась некоторая диссоциация r_2^s- и r_2-субъединиц. Поэтому, если они инкубируются вместе до добавления c_3-субъединиц, картина электрофоретического разделения гибридов АТСазы будет размазанной, она будет представлять собой семь неразрешаемых полос. Если же вместо этого быстро смешать между собой r_2^s-, r_2- и c_3-субъединицы, реассоциированный фермент проявится в виде четырех дискретных электрофоретических полос. Им соответствуют соединения c_6r_6, $c_6r_4r_2^s$, $c_6r_2r_4^s$ и $c_6r_6^s$. Это указывает на то, что r_2-элементы сохраняют свою индивидуальность в нативном ферменте, и субъединичный состав последнего можно записать в виде $(c_3)_2(r_2)_3$ (рис. 2.48,*B*). Доказательством того, что два мономера в каждом r_2-элементе контактируют между собой в нативной структуре, может

служить реконструкция активной АТСазы из чистых r_2-димеров, ковалентно сшитых с *бис*-имидоэфиром.

Структура АТСазы, определенная методом рентгеноструктурного анализа, схематически изображена на рис. 2.49. Как мы видим, эта структура прекрасно соответствует данным, полученным для растворов. Две каталитические субъединицы имеют симметрию C_3 и по форме напоминают треугольник. Они не контактируют между собой, что объясняет, почему в отсутствие регуляторных субъединиц не происходит реакция $2c_3 \leftrightarrows (c_3)_2$. Аналогично этому область контактов между отдельными r_2-субъединицами весьма незначительна. Это согласуется с тем обстоятельством, что ие удалось обнаружить реакцию $3r_2 \rightleftarrows (r_2)_3$ в отсутствие каталитических субъединиц. Состоящая из нескольких слоев структура АТСазы (сэндвич) имеет в центре большую полость, заполненную растворителем. Это совершенно новая для нас особенность белковой молекулы. Из структурных данных можно заключить, что, кроме c_3 и r_2, некоторые совокупности субъединиц также должны быть стабильны или по крайней мере метастабильны. В самом деле, ассоциат $c_3(r_2)_3$ можно приготовить, добавляя небольшие количества c_3 к имеющимся в избытке r_2-субъединицам. Хитрость здесь состоит в том, чтобы быстро «собрать» все субассоциаты c_3 и не допустить осуществления более предпочтительной реакции $c_3 + c_3(r_2)_3 \rightarrow (c_3)_2(r_2)_3$. Когда небольшое количество r_2 добавляется к имеющимся в избытке c_3, образуются субагрегаты $(c_3)_2(r_2)_2$. Они, по-видимому, имеют структуру «сэндвича», у которого одна треть регуляторной части «откушена». Чистый ассоциат $(c_3)_2(r_2)_2$ быстро переходит в нативную АТСазу при добавлении соответствующего количества r_2.

СТАБИЛЬНОСТЬ ЧЕТВЕРТИЧНЫХ ГЛОБУЛЯРНЫХ СТРУКТУР

Из рассмотренных нами примеров АТСазы и других объектов совершенно очевидно, что для понимания свойств, определяющих стабильность четвертичной структуры, надо изучить контакты между субъединицами. Число четвертичных структур, определенных с высоким разрешением методом рентгеноструктурного анализа, пока довольно ограниченно. Однако весьма интересно было бы узнать, что может служить ключом к разгадке сил, связывающих субъединицы между собой. В табл. 2.9 суммированы данные о типах взаимодействий между субъединицами, находящимися в контакте. Как видно из таблицы, всегда преобладают вандерваальсовы контакты. Это говорит о том, что между ассоциированными субъединицами существует хорошее стерическое соответствие. Выше для отдельных субъединиц белка была рассчитана плотность упаковки внутренних остатков. Аналогичный подход можно использовать для расчета плотностей межсубъединичных контактов. Результаты показали, что плотность упаковки боковых цепей на поверхностях контактирующих субъединиц почти такая же, как и в кристаллах аминокислот. Другими словами, поверхности белковых молекул в значительной степени комплементарны друг другу.

Отдельный вандерваальсов контакт «сто́ит» только несколько сотен калорий. Поскольку на поверхности субъединицы насчитывается до 100 и более контактов, они могут внести суммарный вклад в энергию ассоциации субъединиц от 10 до 50 ккал. Однако для того, чтобы соединить две субъединицы между собой, необходимо удалить имеющийся на их поверхности растворитель. Это в свою очередь вызывает потерю большого числа вандерваальсовых контактов. Таким образом, чистая вандерваальсова энергия ассоциации субъединиц слишком мала для того, чтобы полностью объяснять взаимодействие субъединиц.

Константа диссоциации для системы, состоящей из двух субъединиц, лежит в области от 10^{-8} до 10^{-16} моль · $л^{-1}$. Это соответствует свободной энергии димеризации, варьирующей от − 11 до − 22ккал/моль при 25°С. При связывании двух белковых субъединиц с образованием комплекса теряются три поступательные и три вращательные степени свободы, поскольку две субъединицы уже не могут двигаться независимо друг от друга. Многие боковые цепи, обладающие достаточной свободой перемещения на поверхности изолированной молекулы белка, оказываются неподвижными на поверхности субъединиц. Эти эффекты приводят к уменьшению энтропии, что должно компенсироваться энергией взаимодействия двух субъединиц, если образовавшийся комплекс стабилен. Оценки таких потерь энтропии в растворе не особенно надежны, но известно, что они варьируют от 70 до 100кал · $град^{-1}$ · $моль^{-1}$, что соответствует свободной энергии 20-30ккал при комнатной температуре. Следовательно, для компенсации потери энтропии и получения при этом отрицательных значений ΔG^0 для димеризации энергия взаимодействий остатков двух субъединиц должна варьировать от − 30 до − 50ккал · $моль^{-1}$.

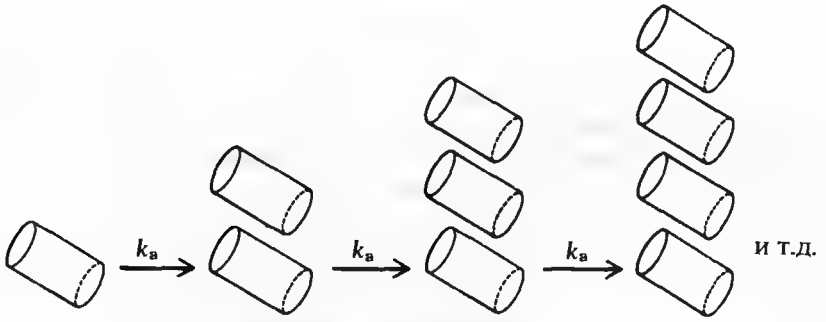
Значительная часть этой энергии, по-видимому, обусловлена гидрофобными взаимодействиями, обсуждаемыми в гл. 5. Было подсчитано, что при тесном контакте двух белковых субъединиц от 1000 до 2000 $Å^2$ площади поверхности, ранее доступной для растворителя, оказывается закрытой. Если эти комплементарные поверхности состоят в основном из неполярных гидрофобных остатков, то прекращение их контактирования с водой ведет к значительному выигрышу в энергии. По некоторым оценкам уменьшение площади контакта с водой на 1 $Å^2$ дает выигрыш энергии 25 кал · $моль^{-1}$. Это приводит к тому, что вклад гидрофобных взаимодействий составляет от − 25 до − 50 ккал · $моль^{-1}$ на две белковые субъединицы. Таким образом, даже без детального рассмотрения специфических полярных контактов между двумя субъединицами можно качественно объяснить энергию их ассоциации в основном гидрофобными эффектами.

Важную роль в образовании четвертичной структуры иногда играют элементы вторичной структуры. В алкогольдегидрогеназе (см. рис. 2.24), конканавалине А и инсулине образование пар субъединиц ведет к формированию протяженных β-слоев. Как видно из табл. 2.9, в двух последних белках обнаружено довольно много водородных связей на поверхности контакта. Слои могут использоваться в областях контакта и иначе. В конканавалине А и глицеральдегид-3-фосфатдегидрогеназе одна поверхность слоя в изолированной субъединице находится на поверхности белка. Когда субъединицы соединяются вместе, их β-слои, плотно укладываясь друг на друга, образуют компактный массив, в котором два слоя связаны осью симметрии второго порядка.

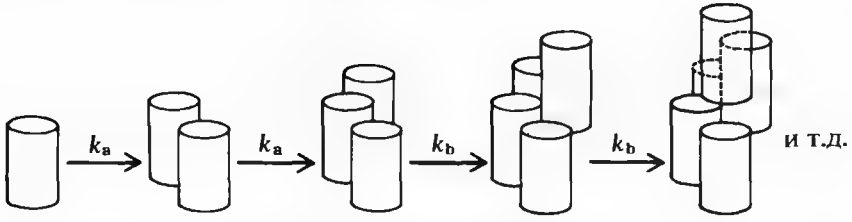
СПИРАЛЬНЫЕ ЧЕТВЕРТИЧНЫЕ СТРУКТУРЫ

До сих пор мы рассматривали белки с относительно простой четвертичной структурой. Большинство из приведенных выше положений справедливы и для более сложных агрегатов, таких как головки вирусов или спиральные образования из субъединиц, найденные в актиновых полимерах, микротрубочках, отростках из бактериофагов и еще более сложных структурах вроде поперечнополосатых мышц и волокон коллагена. Однако для таких образований из субъединиц необходимо учитывать несколько новых факторов. Для их оценки полезно начать с рассмотрения термодинамических аспектов образования линейных слоев субъединиц.

Допустим, что k_a — константа равновесия реакции присоединения одного мономера к концу цепи — не зависит от длины последней, как схематически показано на рис. 2.50.



Линейная полимеризация



Спиральная полимеризация

РИС.2.50. Два основных механизма полимеризации субъединиц. На рис.2.51 схематически изображено предполагаемое поведение каждого из них в зависимости от концентрации белка. [F. Oosawa, M. Kasai, J. Mol. Biol., **4**, 12 (1962).]

Тогда можно не учитывать непосредственные взаимодействия олигомеров или полимеров[1]. Относительные молярные концентрации M цепей как функции длины можно выразить просто через концентрации m свободного мономера:

$$\begin{aligned}(M_2) &= k_a m^2 \\ (M_3) &= k_a m (M_2) = k_a^2 m^3 \\ (M_i) &= k_a m (M_{i-1}) = k_a^{i-1} m^i\end{aligned} \qquad (2.5)$$

Суммарная концентрация полимерных цепей (с длинами, равными 1 и большему числу) равна

$$C_{\text{сумм}} = \sum_{i=1}^{\infty} (M_i) = \sum_{i=1}^{\infty} k_a^{i-1} m^i = (1/k_a) \sum_{i=1}^{\infty} (k_a m)^i = m/(1 - k_a m) \qquad (2.6)$$

где использовалось то обстоятельство, что при $0 \leqslant x \leqslant 1$ (т.е. $m < k_a^{-1}$) для бесконечной геометрической прогрессии $\sum x^i = x/(1 - x)$.

---

[1] Например, образование тетрамера из двух димеров должно иметь ту же константу равновесия k_a, что и образование димера из двух мономеров: $(M_4) = k_a (M_2)^2 = k_a^3 m^4$, что совпадает с уравнением (2.5).

Суммарное количество мономера, присутствующее в растворе, можно получить суммированием числа мономеров, содержащихся в цепях разной длины:

$$m_{\text{сумм}} = \sum_{i=1}^{\infty} i\,(M_i) = \sum_{i=1}^{\infty} i k_a^{i-1} m^i = m \sum_{i=1}^{\infty} i k_a^{i-1} m^{i-1} \tag{2.7}$$

Суммирование в (2.7) легко выполняется, если учесть, что каждый член этой суммы есть $[d/d(k_a m)](k_a m)^i$. Поскольку операции суммирования и дифференцирования коммутируют между собой, уравнение (2.7) принимает вид

$$m_{\text{сумм}} = m[d/d(k_a m)]\left[\sum_{i=1}^{\infty} (k_a m)^i\right] \tag{2.8}$$

Мы уже вычислили эту сумму в уравнении (2.6). Поэтому

$$m_{\text{сумм}} = m[d/d(k_a m)][k_a m/(1 - k_a m)] = m/(1 - k_a m)^2 \tag{2.9}$$

$m_{\text{сумм}}$ обычно известно, и поэтому определение одного из параметров m или k_a дает возможность вычислить другой. Затем, используя уравнение (2.5), легко рассчитать концентрации для всех длин.

Среднечисленное значение степени полимеризации равно среднему числу мономеров, приходящемуся на одну цепь:

$$\langle i \rangle = m_{\text{сумм}}/C_{\text{сумм}} = 1/(1 - k_a m) \tag{2.10}$$

Нас интересует изменение величины $\langle i \rangle$ в зависимости от $m_{\text{сумм}}$ — единственной переменной, которая непосредственно регулируется экспериментатором. По мере постепенного увеличения $m_{\text{сумм}}$ от нулевого значения величина m [как следует из уравнения (2.9)] медленно возрастает, и так же ведет себя $\langle i \rangle$ [см. уравнение (2.10) и рис. 2.51, *А*]. При любом данном значении $\langle i \rangle$ длины цепей могут быть распределены в широком интервале значений.

Рассмотрим теперь простую модель спиральной полимеризации, представленную на рис. 2.50. На нем изображена спираль, содержащая три мономера на один виток, однако все последующие соображения справедливы для любой спирали. До тех пор, пока первый виток спирали не завершен, каждое прибавление мономера может быть описано одним и тем же значением константы ассоциации k_a. Однако при начале заполнения второго витка каждый новый мономер образует контакты уже с двумя субъединицами, а не с одной. Поэтому свободная энергия такой ассоциации субъединиц изменится и процесс добавления новых субъединиц будет характеризоваться новой константой k_b. Поскольку при формировании структуры контакты между субъединицами энергетически выгодны, то $k_b \gg k_a$. Интуитивно можно предположить, что данная модель приведет к представлению о кооперативной полимеризации. Цепи, длина которых больше одного витка спирали, могли бы расти за счет коротких цепей, поскольку для них константы связывания новых мономеров выше.

Пусть j — длина самой короткой спирали с прочным связыванием мономеров. Тогда равновесие реакции полимеризации будет выражаться следующими уравнениями:

$$\begin{aligned} (M_i) &= k_a^{i-1} m^i \quad \text{для} \quad i < j \\ (M_i) &= k_a^{j-2} k_b^{i-j+1} m^i \quad \text{для } i \geqslant j \end{aligned} \tag{2.11}$$

В примере, показанном на рис. 2.50, $j = 4$. Как и в случае линейной полимеризации, полезно рассчитать суммарную концентрацию мономеров белка.

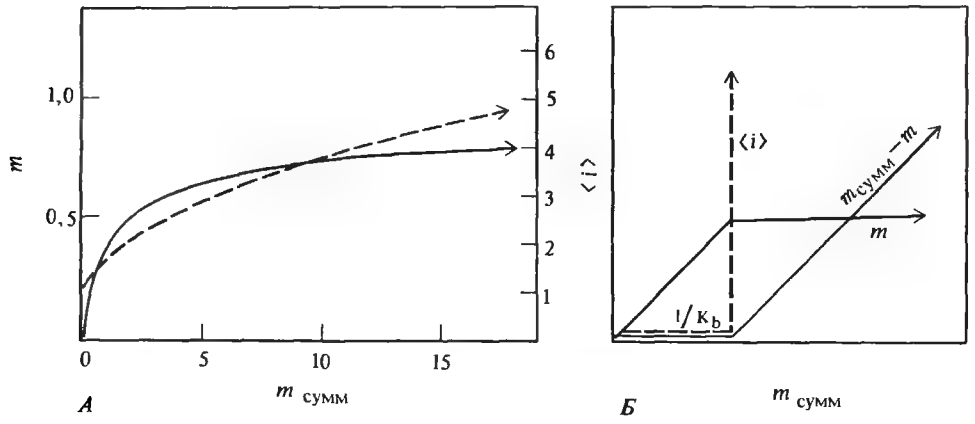


РИС.2.51. Физические особенности двух механизмов полимеризации, показанных на рис. 2.50. *А*. Линейная полимеризация. Концентрация свободного мономера m (сплошная линия) и среднечисленное значение степени полимеризации $\langle i \rangle$ (пунктирная линия) монотонно возрастают с ростом суммарной концентрации белка $m_{\text{сумм}}$. *Б*. Спиральная полимеризация. Концентрация свободных мономеров m (сплошная кривая) линейно растет до критического значения концентрации и далее остается постоянной. До тех пор пока концентрация $m_{\text{сумм}}$ не достигает критического значения, среднечисленное значение степени полимеризации (пунктирная кривая) равно 1, а затем резко увеличивается. Количество вещества в длинных полимерах (цветная кривая) равно нулю до достижения критической концентрации, а затем линейно увеличивается с ростом суммарной концентрации. [F. Oosawa, M. Kasai, J. Mol. Biol., **4**, 12(1962).]

$$m_{\text{сумм}} = \sum_{i=1}^{\infty} i\,(M_i) = \sum_{i=1}^{j-1} i k_a^{i-1} m^i + k_a^{j-2} \sum_{i=j}^{\infty} i k_b^{i-j+1} m^i \qquad (2.12)$$

Второй член этого уравнения можно представить в удобной форме, если его умножить и разделить на k_b^{j-2}:

$$(k_a^{j-2}/k_b^{j-2}) \sum_{i=j}^{\infty} i k_b^{i-1} m^i = (k_a/k_b)^{j-2} \left(\sum_{i=1}^{\infty} i k_b^{i-1} m^i - \sum_{i=1}^{j-1} i k_b^{i-1} m^i \right) \qquad (2.13)$$

Первая из двух сумм в правой части равенства может быть вычислена, как в уравнении (2.8). Это дает возможность представить суммарную концентрацию мономера в виде

$$m_{\text{сумм}} = \sum_{i=1}^{j-1} i k_a^{i-1}\, m^i + \sigma \left([m/(1 - k_b m)^2] - \sum_{i=1}^{j-1} i k_b^{i-1} m^i \right) \qquad (2.14)$$

где $\sigma = (k_a/k_b)^{j-2}$ много меньше единицы.

Теперь посмотрим, что происходит при постепенном увеличении от нуля суммарной концентрации мономеров в растворе. Вначале $mk_a \ll mk_b \ll 1$, так что в уравнении (2.14) единственным большим членом будет первый член в первой сумме. То, что $m_{\text{сумм}}$ приблизительно равно m, означает, что все добавленные белковые молекулы остаются мономерами. С ростом $m_{\text{сумм}}$ до значения $1/k_b$ величина m увеличивается до тех пор, пока mk_b не станет близким к единице. Отметим, что при этом величина mk_a все еще остается значи-

тельно меньше 1 и поэтому концентрация олигомеров длиной, меньше чем j, остается малой. Член $(1 - k_b m)^{-2}$ становится очень большим по мере того, как m приближается к $1/k_b$. Так как $m_{\text{сумм}}$ — величина конечная, концентрация свободных мономеров никогда не может достичь величины $1/k_b$, однако при малых значениях σ она может приблизиться к этой величине. При $m_{\text{сумм}}$, бóльших $1/k_b$, концентрация свободных мономеров остается постоянной и близкой к $1/k_b$, а избыточные молекулы белка будут включаться в спиральные полимеры (см. рис. 2.51, *Б*).

Существенная особенность спиральной полимеризации заключается в том, что существует критическая концентрация белка, приблизительно выражаемая равенством $m_{\text{сумм}} = k_b^{-1}$. При концентрациях, меньших критической, все белки существуют в виде мономеров. При концентрациях, бóльших критической, содержание мономеров в растворе сохраняется постоянным, а весь дополнительный белок объединяется в довольно длинные полимеры. Численное значение средней степени полимеризации дается приближенным равенством $\langle i \rangle \approx 1/(1 - k_b m)$, по аналогии с уравнением (2.8). Однако, поскольку $k_b m \approx 1$, величина $\langle i \rangle$ будет огромной. О физических причинах этого уже говорилось. Как только достигается критическая длина j, дальнейшее присоединение мономеров к этому ассоциату становится более выгодным, чем создание нового, поскольку $k_b \gg k_a$. Следовательно, при спиральной полимеризации можно ожидать существования равновесия между длинными полимерами и свободными мономерами. Именно это наблюдалось в системах типа актина и микротрубочек.

ОПРЕДЕЛЕНИЕ ДЛИНЫ СПИРАЛЬНЫХ АССОЦИАТОВ

Описанный выше механизм спиральной полимеризации представляет собой простой способ образования больших агрегатов. И другие механизмы полимеризации, в которых вслед за стадией, неблагоприятной для инициации полимеризации, следуют благоприятные стадии быстрого роста агрегатов, могут приводить к такому же результату. Ни один из этих механизмов не дает узкого распределения полимеров по длинам, однако in vivo наблюдается узкое распределение для таких структур, как палочкообразные вирусы и хвостовые отростки бактериофагов. Эти структуры можно воссоздавать in vitro из их компонентов, причем кривая распределения не меняется. Поэтому, чтобы объяснить образование такой структуры, надо принимать во внимание только те макромолекулярные компоненты, которые обнаруживаются в уже полностью агрегированной структуре.

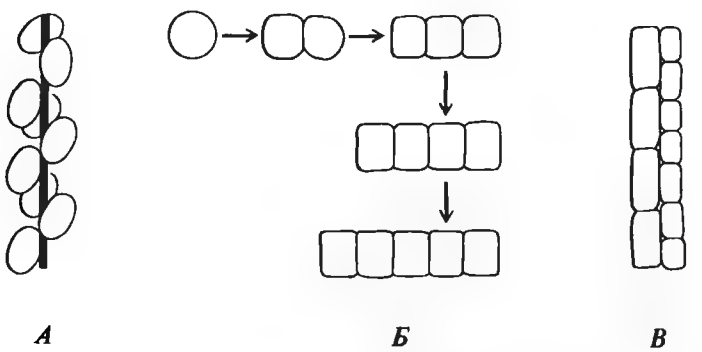


РИС.2.52. Схематическое представление механизмов образования узкого распределения полимерных агрегатов по длинам (объяснение в тексте). *А*. Белок или другая молекула играет роль ядра агрегата. *Б*. Модель с постепенным увеличением напряжения. *В*. Сборка по принципу коннуса.

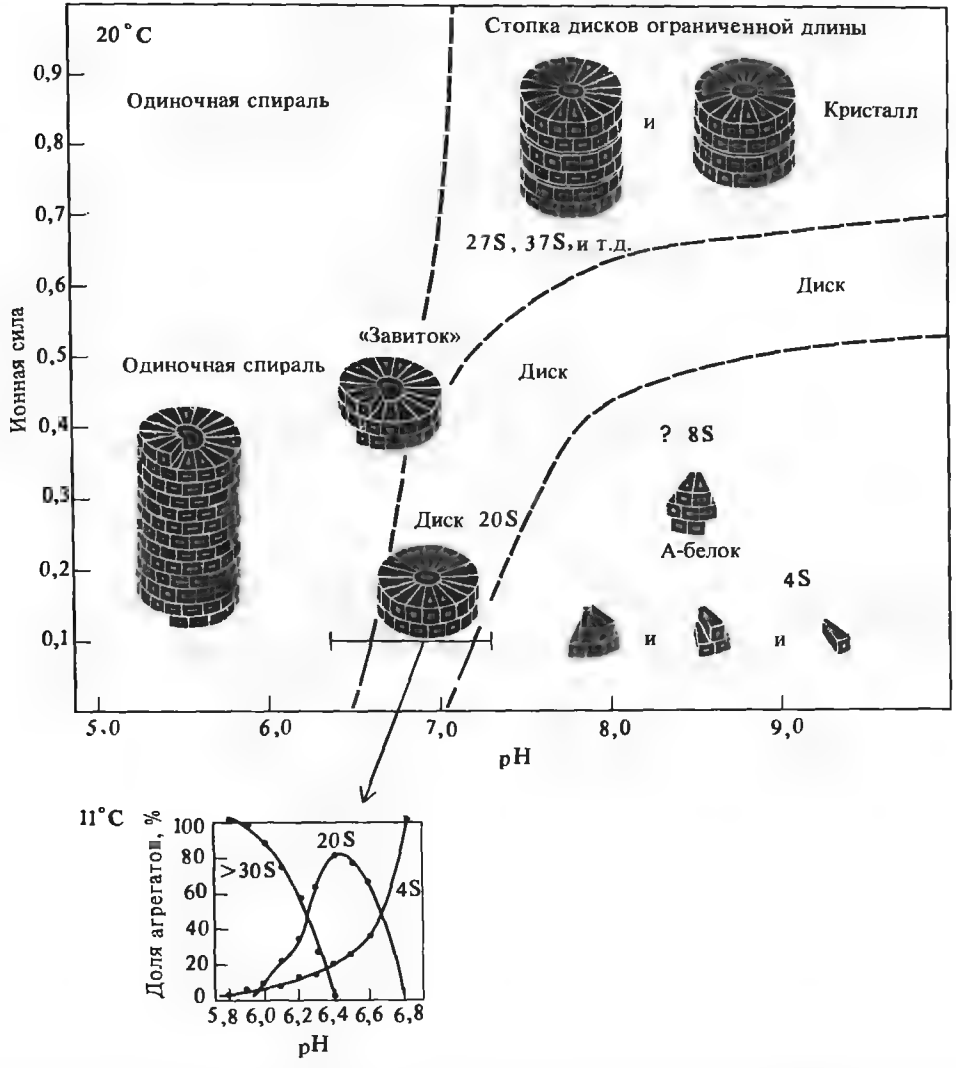


РИС.2.53. Фазовая диаграмма самосборки белковых агрегатов вируса табачной мозаики при 20 °С в зависимости от pH и ионной силы раствора. Пунктирными линиями выделены области, в которых можно обнаружить большие агрегаты, но не обязательно отсутствуют малые. Здесь можно усмотреть аналогию с критическими концентрацнями. Находящийся на границе фазовых областей агрегат типа «завиток» является метастабильиым. На рисунке показан также результат типичного эксперимента, в котором pH изменяется при постоянной ионной силе, равной 0,1. Агрегаты с константой седиментации, большей 30 S, — спирали с широким распределением по длинам. [A. Klug, A. Durham, Cold Spring Harbor Symp. Quant. Biol., vol. 36 (1971).]

Подобное узкое распределение по длинам можно объяснить с помощью нескольких механизмов полимеризации. Некоторые примеры представлены на рис. 2.52.

В матричном механизме отдельная длинная молекула служит матрицей для связывания и упорядочивания белков. Ее длина, таким образом, совпадает с длиной конечной структуры. Этот механизм хорошо объясняет образование многих палочкообразных вирусов типа вируса табачной мозаики. В этом случае каждый из трех нуклеотидов одноцепочечной молекулы РНК вируса служит для связывания одной субъединицы белка оболочки. Если сборка белковой оболочки вируса происходит в отсутствие РНК, то образуются палочкообразные структуры различной длины, во всем остальном очень похожие на оболочки обычных вирусов. В зависимости от концентрации белка и условий, характеризующих окружение вируса, средние значения длин могут быть как больше, так и меньше, чем у обычного вируса, хотя в каждом случае в отсутствие РНК наблюдается широкое распределение по длинам.

При более подробном изучении самосборки белковой оболочки вируса табачной мозаики можно найти условия, в которых помимо мономеров и длинных спиралей существуют несколько промежуточных форм (рис. 2.53). Соответствующий рисунок по аналогии с описанием макроскопических фаз чистых веществ или простых смесей можно назвать фазовой диаграммой. Поскольку переходы между различными четвертичными структурами при изменении окружающей среды происходят довольно резко, каждую форму можно уподобить макроскопической фазе. Из рис. 2.53 видно, что модель простой спиральной полимеризации не может адекватно описать процесс сборки вируса табачной мозаики. Более реальная модель должна включать дополнительные параметры для учета ряда промежуточных состояний. И в любом случае любая плодотворная модель должна учитывать принципы образования зародышей и роста цепей для объяснения появления очень длинных стерженьков.

Заслуживают краткого упоминания еще две модели, приводящие к дискретным значениям длины полимера. В первой из них постулируется, что для присоединения первого мономера существуют весьма благоприятные условия, тогда как присоединение каждой последующей субъединицы становится все менее и менее выгодным. Например, константы ассоциации для реакции $M_1 + M_i \rightleftarrows M_{i+1}$ могут иметь вид $k_i = kx^i$, где $x < 1$. Эту модель часто называют моделью с постепенным увеличением напряжения. Свободная энергия напряжения на один субъединичный контакт составляет $-RT\ln x$. Эта энергия может или распределяться между всеми субъединицами (приводя в результате к однородной структуре), или периодически накапливаться в определенных местах. Если увеличение напряжения при агрегировании субъединиц достигнет большой величины, то размеры цепи будут определяться тем числом субъединиц, при котором связывание дополнительных мономеров становится энергетически невыгодным. Это приводит к довольно острому пику в кривой распределения длин полимеров. Данный механизм может быть ответствен за распределение длин хвостовых отростков некоторых бактериофагов, для которых не доказано наличие молекулы-матрицы.

В основе второй модели (см. рис. 2.52,*В*) лежит принцип нониуса. Если пытаться сгруппировать молекулы разной длины, то только определенные соотношения их длин приведут к структуре, которая не будет иметь выступающих концов. Такая модель требует присутствия в агрегате субъединиц нескольких типов — это могут быть различные типы белков или различные третичные структуры одного белка. И хотя до сих пор не обнаружены системы, устроенные по принципу нониуса, сама гипотеза представляется интересной и ее нельзя упускать из виду, обсуждая более сложные биологические ассоциаты.

В живой клетке при формировании линейных или спиральных ассоциатов могут существовать и более сложные механизмы, определяющие их длину. В клетке могут существовать специфические области, где начинается или заканчивается полимеризация. К тому же внутриклеточная среда вовсе не обязательно должна быть однородной; это может способствовать полимеризации в одних областях и подавлению ее в других. Очевидно,

что некоторые полимеризованные структуры, такие как микротрубочки и волокна актина, простираются на заметные расстояния внутри клеток. Они, по-видимому, играют жизненно важную роль в определении формы клетки, ее передвижения и деления, а также, возможно, в организации связи между поверхностью клетки и ее внутренними областями. Факторы, регулирующие формирование таких структур или их размер, чрезвычайно важны для понимания биологии клетки.

Краткие выводы

Белки синтезируются из 20 аминокислот, но иногда содержат и другие химические группы, например ионы металлов, органические простетические группы и сахара. Аминокислоты можно грубо разделить на два класса: полярные и неполярные.

Различные белки значительно отличаются друг от друга по своему аминокислотному составу, и существует очевидная корреляция между составом и структурой или функцией. Аминокислотные последовательности определить относительно легко, и часто эти данные позволяют получить дополнительную информацию о структуре и функции. Более того, знание последовательности практически необходимо для определения полной структуры белка.

В белках наиболее распространены три типа вторичных структур: α-спирали, β-слои и β-изгибы. Все вместе они содержат более половины всех аминокислот белковой молекулы. В тех белках, в которых часто встречаются остатки пролина, α-спирали и β-слои образуются редко; вместо них могут формироваться спирали полипролинового типа.

Третичные структуры большинства белков организованы в плотно упакованные глобулы или в несколько глобулярных доменов. В таких свернутых структурах неполярные боковые группы стремятся оказаться внутри молекулы, а боковые группы полярных остатков — снаружи. Между аминокислотными остатками, а также между остатками и растворителем, по-видимому, образуется обширная сеть водородных связей. Сравнение укладки цепи в различных белках дает возможность глубже проникнуть в их структуру. Семейства белков со сходными или почти тождественными структурами встречаются довольно часто, но особенно удивительно сходство типов организации β-слоев и α-спиралей в белках с совершенно различными функциями.

Обычно четвертичные структуры делят на два типа. Для многих белков типично глобулярное расположение субъединиц. В тех случаях, когда субъединицы тождественны, они почти всегда расположены симметрично; однако разрешены только некоторые определенные четвертичные структуры, так как сами по себе субъединицы асимметричны. Возможны следующие типы симметрии: аксиальная, диэдрическая и кубическая. Другие белки объединяются в спиральные четвертичные структуры, в которых соседние субъединицы связаны между собой винтовой симметрией. Длину таких структур или распределение длин регулируют различные механизмы.

ЗАДАЧИ

2.1. Из трех источников — трубкозуба (А), бобра (В) и верблюда (С) — выделены белки, эквивалентные по своим функциям и сходные по структуре. Электрофоретические подвижности трех белков несколько отличаются друг от друга, но для каждого существует только одна полоса. Если смешать А и В или А и С, то на электрофореграмме наблюдаются три полосы; если смешать В и С — наблюдается 5 полос. Объясните этот факт.

2.2. Дайте простое объяснение, почему тример и диск являются относительно стабильными олигомерными формами белка вируса табачной мозаики (см. рис. 2.53).

2.3. Допустим, что белок состоит из одной полипептидной цепи и из данных рентгеноструктурного анализа известно, что отдельная субъединица обладает почти точной симметрией C_2. Какой вывод можно сделать о первичной структуре и о расположении N- и С-концов в третичной структуре?

2.4. Установлено, что рибосомный белок, состоящий из 85 аминокислот, имеет в рибосоме длину 150 Å. В растворе он существует в виде изолированной субъединицы. Что можно сказать о возможной вторичной структуре этого белка в растворе и в том случае, когда он находится в контакте с другими белками в рибосоме?

2.5. Последовательность аминокислот в гипотетическом белке записывается в виде

N-Ala-Arg-Val-Ser-Met-Lys-Ilu-Glu-Ala-Lys-Gly-Asp-Trp-Thr-Gly-Gly-Gln-Met-Thr-Gly-Asp-Ala-Asn-Phe-Arg-Ala-Ser-Val-Glu-Leu-C.

Установлено, что этот белок образует очень стабильную вторичную структуру и существует в растворе в виде димера. Изолированный мономер легко связывается с мембранами, димер — не связывается. Какие можно сделать заключения о вторичной, третичной и четвертичной структуре белка?

ЛИТЕРАТУРА

Общая

Boyer P. D., ed., 1970. The Enzymes, 3d ed., vol. 1, New York, Academic Press. [Особенно рекомендуются следующие статьи. Reed L. J., Cox D. J. Multienzyme complexes. p. 213; Smith E. L. Evolution of enzymes, p. 267; Koshland D. E. The molecular basis of enzyme regulation, p. 342; Stadtman E.A. Mechanisms of enzyme regulation in metabolism, p. 395; Atkinson D. F. Enzymes as control elements in metabolic regulation, p. 461.]

Cold Spring Harbor Laboratory, 1972. Structure and function of proteins at the three-dimensional level, Cold Spring Harbor Symposia in Quantitative Biology. vol. 36.

Dickerson R., Geis I., 1969. The Structure and Action of Proteins, New York, Harper and Row. [Ясно изложенное и хорошо иллюстрированное введение в проблему.]

Gutfreund H., ed., 1974. Chemistry of Macromolecules, London, Butterworths (Biochemistry Series One, vol. 1). [Превосходный набор обзорных статей. Особенно рекомендуются следующие: Williams J. The primary structure of proteins in relation to evolution, p. 1; Muirhead H. The three-dimensional structure of proteins, p. 57; Joynson M. A. The subunit structure of proteins, p. 85; Knowles J. P. The functions of proteins as devices, p.149.]

Haschemeyer R. H., Haschemeyer A. E. V., 1973. Proteins: A Guide to Study by Physical and Chemical Methods, New York, Wiley. [Охватывает много различных методов и явлений.]

Neurath H., Hill R. C, eds., 1975. The Proteins, 3d ed., New York, Academic Press. [Многотомный курс, содержащий большое количество прекрасных статей.]

Schulz G. E., Schirmer R. H., 1979. Principles of Protein Structure, New York, Springer Verlag. [Текст рассчитан на специалистов, содержит исчерпывающую библиографию.]

Специальная

Chou. P. Y., Fasman G. D., 1978. Empirical predictions of protein conformation, Ann. Rev. Biochem., 47, 251.

Dayhoff M. O., ed., 1972 — 1976. Atlas of Protein Sequence and Structure, Silver Spring, Md., National Biomedical Research Foundation.

Fersht A., 1977. Enzyme Structure and Mechanism, Reading, U. K., W. H. Freeman and Company. [Имеется перевод: Фёршт Э. Структура и механизм действия ферментов. — М.: Мир, 1980.]

Fessler J. H., Fessler L. I., 1978. Biosynthesis of procollagen, Ann. Rev. Biochem., 47, 129.

Oosawa F., 1975. Thermodynamics of the Polymerization of Protein, New York, Academic Press.

Perutz M. F., Lehmann H., 1968. Molecular pathology of human hemoglobin, Nature, 219, 902.

Richardson J. S., 1976. Handedness of crossover connection on βsheets, Proc. Natl. Acad. Sci. USA, 73, 2619.

Rossmann M. G., Argos P., 1977. Taxonomy of protein structure, J. Mol. Biol., 109, 99.

Schachman H. K., 1974. Anatomy and physiology of a regulatory enzyme, aspartate transcarbamylase, The Harvey Lectures, 68, 67.

Глава 3 Структура нуклеиновых кислот

3.1. Свойства нуклеозидов и нуклеотидов

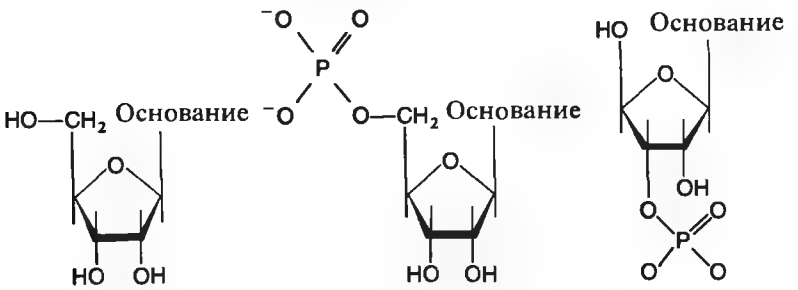
Мономерные единицы нуклеиновых кислот называются нуклеотидами. Они состоят из фосфата, сахара и пуринового или пиримидинового основания (на рис. 3.1 приведены структурные формулы и названия наиболее распространенных оснований и сахаров).

Все обнаруженные в нуклеиновых кислотах сахара относятся к D-стереоизомерам, изображенным на рис. 3.1. Обычные основания и большая часть минорных оснований имеют плоскую структуру, и у них нет стереоизомеров. Лишь несколько минорных оснований (таких, как основание Y) имеют асимметричный атом углерода в боковой цепи. Часть мономерной единицы, содержащая только сахар и основание, называется нуклеозидом, так что нуклеотиды есть просто фосфорные эфиры нуклеозидов. Нуклеозиды обозначают двумя буквами латинского алфавита. Большая буква отвечает основанию, а маленькая — типу сахара. Так, аденозин (который состоит из сахара рибозы и основания аденина) обозначают rA, а дезоксигуанозин (с сахаром 2′-дезоксирибозой и основанием гуанином) — dG. Если из контекста ясно, какой тип сахара имеется в виду, приставку r обычно опускают. При обозначении нуклеотидов к большой букве, отвечающей нуклеозиду, добавляют букву p, соответствующую фосфату (см. выше).

Полинуклеотидную последовательность сокращенно записывают в виде

$$\ldots \text{pNpNpNpNpNpNp} \ldots \text{ или } (\text{pN})_n, \text{ или } (\text{Np})_n$$

Тем самым подчеркивается, что остов полимера построен из остатков сахаров, связанных фосфодиэфирными мостиками (рис. 1.2). В РНК сахар представлен только рибозой и преобладают четыре нуклеозида — A, U, C и G, хотя достаточно часто встречаются и «необычные» нуклеозиды, особенно в транспортных РНК. В ДНК сахар представлен 2′-дезоксирибозой, и практически во всех известных ДНК присутствуют лишь четыре нуклеозида — dA, dT, dC и dG. Исключение составляют ДНК ряда вирусов, в которых один из обычных нуклеозидов замещен частично или полностью минорным нуклеозидом. Например, ДНК некоторых бактериофагов содержит 5-гидроксиметил-dC вместо dC, моно- или дисахариды часто присоединяются к цепи через гидроксил.



Нуклеозид N

5′-нуклеотид pN

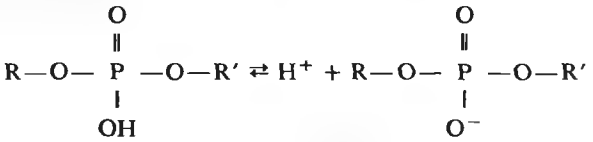
3′-нуклеотид Np

В этой книге, если нет никаких специальных оговорок, мы будем представлять одиночную полинуклеотидную цепь в виде ApBpCp. . . pZ, подразумевая при этом, что каждая фосфодиэфирная связь направлена от 3′- к 5′-гидроксилу сахара, слева направо. Таким образом, 5′-конец молекулы оказывается слева, а 3′-конец — справа. Чтобы было ясно направление антипараллельных цепей, их концы принято отмечать цифрами 5′ и 3′.

ИОНИЗАЦИОННОЕ РАВНОВЕСИЕ НУКЛЕОЗИДОВ И НУКЛЕОТИДОВ

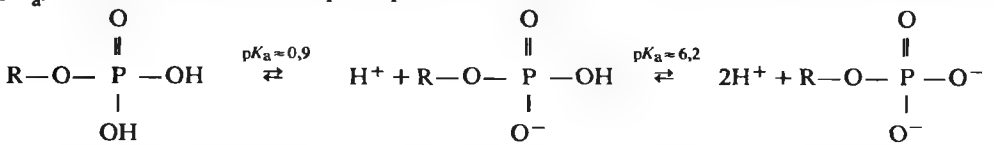
Все обычные нуклеозиды при нейтральных pH незаряжены, что соответствует форме, представленной на рис. 3.1. Большинство необычных нуклеозидов также незаряжены, хотя 7-метил-G при pH 7 заряжен положительно. Напротив, все фосфаты в нуклеотидах или полинуклеотидах — при тех pH, при которых обычно работают с этими соединениями, — заряжены отрицательно.

Фосфодиэфир является сильной кислотой. Значение pK_a единственного способного к ионизации протона близко к 1:



Таким образом, если пренебречь концевыми эффектами, то можно считать, что на каждый остаток полинуклеотидной цепи приходится один отрицательный заряд. Наличием большого суммарного заряда полинуклеотиды существенно отличаются от типичных белков, в которых большинство остатков незаряжено, а положительные и отрицательные боковые цепи обычно представлены в сопоставимом (хотя и неравном) количестве.

Фосфомоноэфирная группа присутствует в изолированных мононуклеотидах или иногда на концах полинуклеотидных цепей. Она имеет два протона, способных к ионизации, с pK_a, типичные значения которых приведены на схеме:



Таким образом, при pH 7 большая часть молекул фосфомоноэфира несет двойной отрицательный заряд. Получаемые при титровании любого мононуклеотида значения pK_a около 1 и 6 несомненно относятся к фосфатной группе, потому что групп с такими pK_a нет в соответствующих нуклеозидах. Более того, эти значения близки к наблюдаемым в простом сахарофосфате. Например, глюкозо-3-фосфат имеет pK_a около 0,8 и 5,7.

У нуклеозидов обычно наблюдается pK_a около 12,4, что существенно отличается от pK_a, характерных для нуклеотидов. Это значение pK_a связано с ионизацией гидроксильной группы сахара (для глюкозы, например, наблюдается $pK_a = 12{,}1$). Отрицательный заряд нуклеотида должен несколько сдвигать pK_a гидроксильной группы сахара в область более высоких pH из-за сильного электростатического отталкивания в троекратно депротонированной отрицательно заряженной форме. Поэтому pK_a гидроксила скорее всего будет находиться в области pH, при которых нуклеотид подвергается гидролизу, и отрыв электрона от сахара в составе нуклеотида или полинуклеотида наблюдать не удастся.

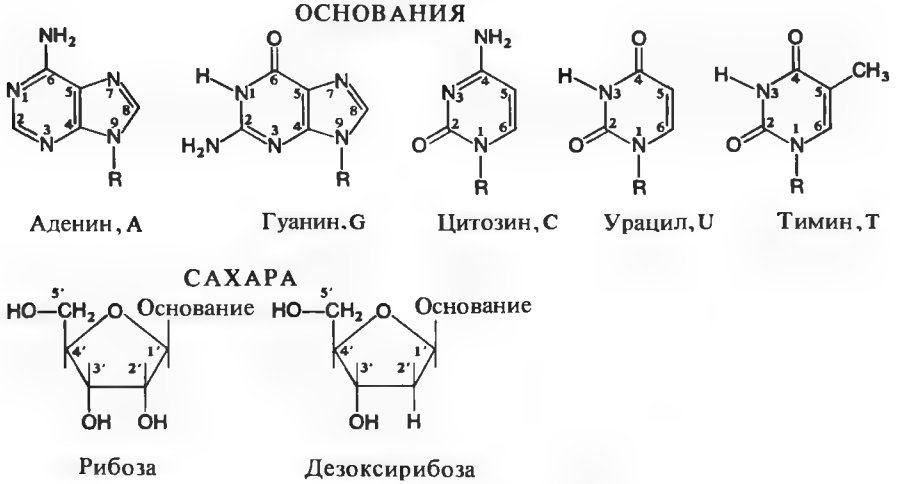
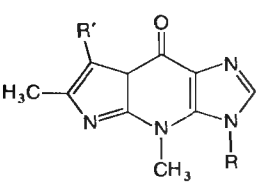


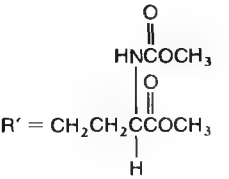
РИС.3.1. Структурные формулы компонентов нуклеиновых кислот. *А*. Обычные основания и сахара. *Б*. Некоторые из минорных оснований и нуклеозидов. Приведены также стандартные сокращения.

Поведение при титровании большинства нуклеотидов и нуклеозидов очень сложно, потому что они содержат много групп, являющихся потенциальными донорами или акцепторами протонов. Прежде чем обсуждать этот вопрос, отметим, что даже для неионизованной формы оснований не сразу очевидно существование единственной локализации протонов на гетероатомах. Каждое основание в принципе может принимать ряд таутомерных форм. Несколько форм, между которыми существует равновесие, показаны на рис. 3.2. Из совокупности доступных кристаллографических и спектроскопических данных можно сделать достаточно надежный вывод, что в водных растворах преобладают кето-амино-таутомеры, показанные на рис. 3.1.

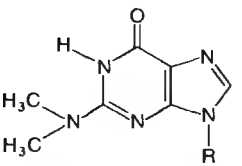
Наиболее эффективными методами исследования таутомерных форм являются ЯМР- и ИК-спектроскопия, поскольку с их помощью часто удается наблюдать спектральные полосы от индивидуальных протонов. Идентифицировать эти полосы можно путем сравнения со спектрами соответствующих производных, в которых возможности таутомеризации ограничены, скажем, в результате замены некоторых протонов метильными группами. По оценкам, разность в энергии между основными таутомерными формами и некоторыми из форм, приведенных на рис. 3.2, невелика, поэтому вопрос о том, играют ли необычные таутомеры какую-нибудь роль в формировании структуры или функционировании нуклеиновых кислот, остается открытым. В настоящее время прямые данные на этот счет отсутствуют.

Если присоединить или удалить протон из основания, то получившийся продукт, как правило, по-прежнему сможет принимать несколько таутомерных форм. Идентификация наиболее стабильных из них эквивалентна ответу на вопрос о том, какие протоны удалены или какие группы связывают протоны при значениях pH, соответствующих наблюдаемым pK_a. С этой целью было выполнено большое число экспериментальных работ; результаты для четырех обычных рибонуклеозидов суммированы на рис. 3.3. Отметим, что основания A и C могут находиться в нейтральной или положительной форме, U — в нейтральной или отрицательной, а G может существовать во всех трех состояниях. Наличие фосфатов сдвигает все значения pK_a, приведенные на рис. 3.3, вверх на несколько десятых

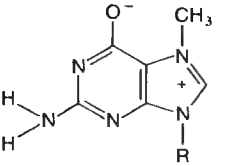




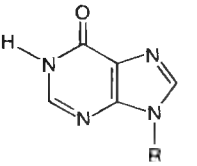
Основание Y



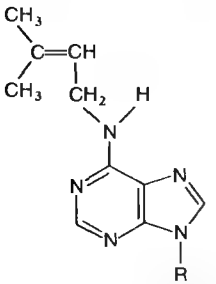
N^2-диметилгуанозин
2dmG



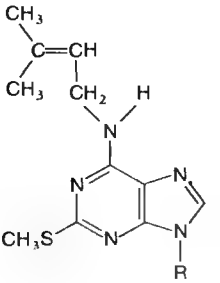
7-метилгуанозин
7mG



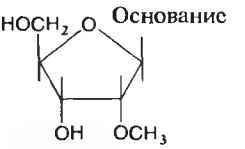
Инозин
I



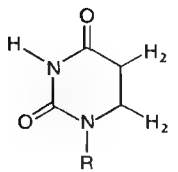
N^6-Δ^2-изопентенил-аденозин
6iA



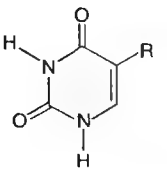
N^6-Δ^2-изопентенил-2-метилтиоаденозин
2ms6iA



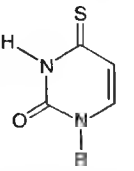
2′O-метилнуклеозид
2′Om (A, C, G или U)



Дигидроуридин
D



Псевдоуридин
ψ



4-тиоуридин
4s

Б

единицы pH. Этот сдвиг обусловлен электростатическим притяжением между фосфатами и любыми положительными формами и отталкиванием между фосфатами и любыми отрицательными формами. Наблюдаемый сдвиг больше для 5′-нуклеотидов, чем для 2′- или 3′-изомеров. Это и понятно, поскольку 5′-положение ближе к возможным местам ионизации оснований.

Все свойства дезоксирибонуклеозидов и дезоксирибонуклеотидов, выявляемые при их титровании, хорошо коррелируют со свойствами соответствующих рибокомпонентов, за тем исключением, что для дезоксирибокомпонентов значения pK_a оказываются, как правило, на несколько десятых выше. Для олигонуклеотидов или полинуклеотидов все наб-

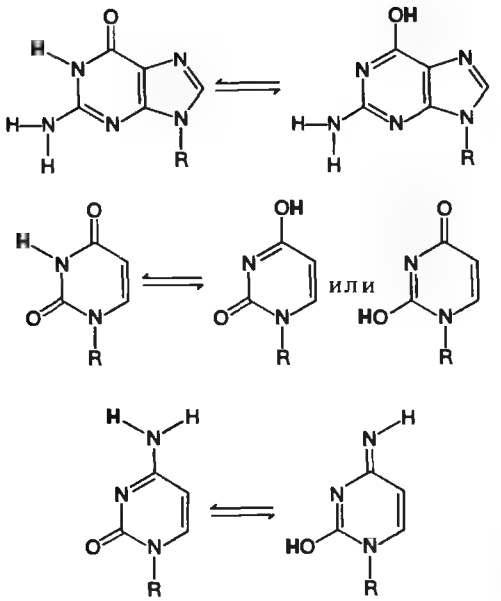


РИС.3.2. Некоторые возможные таутомерные превращения для G, U и C при pH 7.

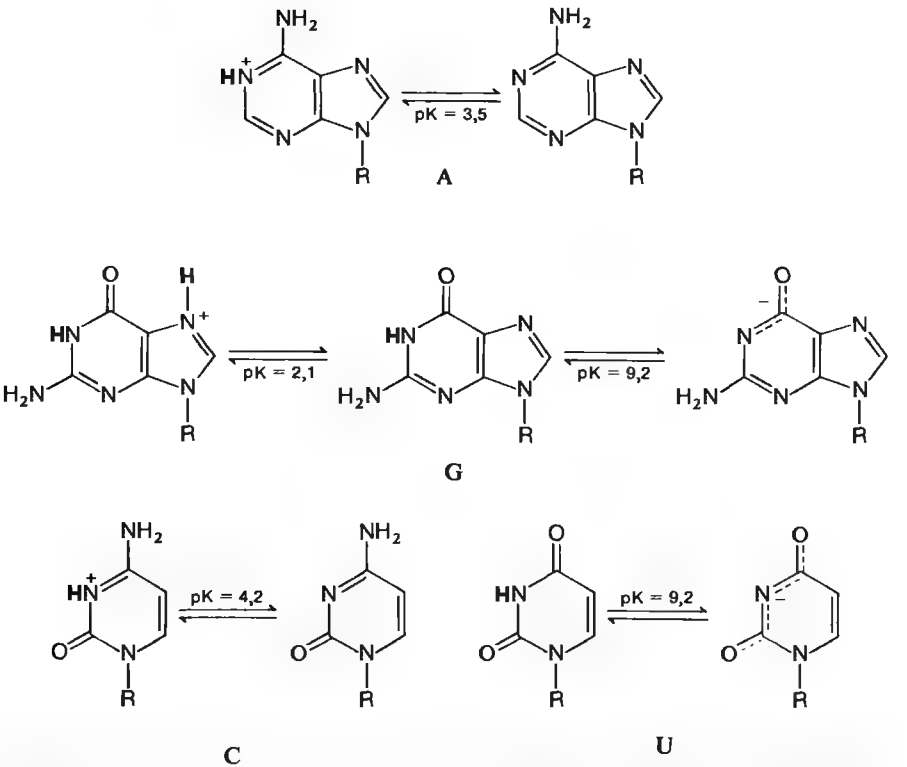


РИС.3.3. Ионизационное равновесие для четырех рибомононуклеозидов. Аналогичные процессы ионизации характерны для дезоксирибомононуклеозидов.

людаемые значения pK_a несколько смещены вследствие электростатического отталкивания между остатками. Более значительные и сложные эффекты наблюдаются, когда группы, способные к ионизации, участвуют в формировании вторичной структуры посредством спаривания оснований. В этих случаях, как и для белков, титрование нуклеотидов оказывается сопряженным с конформационными переходами, что может приводить к очень большим смещениям pK_a. Некоторые примеры такого рода рассмотрены в гл. 22.

Все обычные основания нуклеиновых кислот весьма сходны по своим основным химическим и физическим свойствам. Например, они обладают (или, согласно оценкам, должны обладать) довольно высоким дипольным моментом — от 2,5 до почти 7Д. Таким образом, с чисто электронной точки зрения основания должны рассматриваться как полярные соединения. Несмотря на это, все обычные основания, кроме урацила, почти нерастворимы в воде. Растворимость пуринов в 5-10 раз меньше, чем пиримидинов, но этот диапазон изменения растворимости значительно меньше, чем для разных боковых цепей аминокислот. Таким образом, грубо говоря, можно считать все основания полярными и гидрофобными соединениями (что звучит несколько парадоксально). Во всяком случае, нет никаких оснований классифицировать нуклеотидные остатки подобно тому, как это было сделано для аминокислот в гл. 2.

3.2. Нуклеотидный состав ДНК и РНК

ПРАВИЛА ЧАРГАФФА ДЛЯ ДНК

Подавляющее большинство молекул ДНК являются двухцепочечными, что налагает ограничения на их нуклеотидный состав. Для этих биополимеров всегда $\chi_A = \chi_T$ и $\chi_G = \chi_C$, где χ — мольная доля. Этот факт, открытый и обстоятельно изученный Э. Чаргаффом, был одним из важнейших среди тех данных, которые привели к концепции двойной спирали для ДНК. Он означает, что для описания состава совершенной двухцепочечной ДНК необходима только одна переменная $\chi_{G+C} (= 1 - \chi_{A+T})$. GC-содержание природных ДНК существенно различается. Так, для ДНК некоторых бактерий $\chi_{G+C} = 75\%$. Вероятно, существование столь широкого интервала значений GC-содержания без потери способности организмов хранить необходимую генетическую информацию становится возможным благодаря вырожденности генетического кода и тому факту, что часть ДНК не кодирует никаких белковых последовательностей. У животных $\chi_{G+C} \simeq 40 - 45\%$, и, вообще говоря, ДНК у близких организмов имеют сходное GC-содержание. Выделена очень небольшая минорная фракция эукариотической ДНК с экстремальным нуклеотидным составом — так называемая сателлитная ДНК; для сателлитной ДНК краба, например, χ_{G+C} равна всего 5%. Биологическая роль таких сателлитов недостаточно ясна, но в тех случаях, когда они исследованы более детально, установлено, что обычно эта ДНК состоит из регулярно повторяющихся коротких последовательностей.

В состав некоторых ДНК входят необычные основания, замещающие один (или более) пиримидин или пурин. Например, в ДНК Т-четных фагов, как мы уже упоминали, преобладает глюкозилированный 5-гидроксиметил-С. Он образует с гуанином такие же пары оснований, как обычный цитозин, и правила Чаргаффа остаются справедливыми, если при вычислении χ_C учесть все аналоги цитозина. ДНК фага SP2 содержит dU вместо dT; известно и еще несколько ДНК с необычными пиримидинами. Некоторые основания отдельных ДНК метилированы, причем присутствуют они в очень специальных последовательностях. Эта модификация служит механизмом защиты против внутриклеточных рестриктаз, которые в отсутствие метилирования расщепляли бы ДНК в этих местах.

Для некоторых ДНК, однако, правило Чаргаффа не выполняется. Например, для ДНК бактериофага ∅X174

$$\chi_A = 0{,}25, \quad \chi_T = 0{,}33, \quad \chi_G = 0{,}24, \quad \chi_C = 0{,}18$$

Это означает, что в формировании двухцепочечной структуры, образованной из обычных AT- и GC-пар, могут участвовать максимум $(0{,}18 + 0{,}25) \cdot 2 = 0{,}86$ от общего числа оснований. В действительности ДНК ∅X174 в отличие от ДНК большинства других организмов является одноцепочечной молекулой, и нет никаких оснований считать, что будет происходить достаточно эффективное спаривание оснований в пределах одной цепи. На самом деле доля спаренных оснований значительно меньше 0,86. Выполнение правил Чаргаффа не означает, что ДНК обязательно является двухцепочечной. Способность образовывать такую структуру определяется особенностями нуклеотидной последовательности и может быть выявлена различными физическими методами (гл. 22).

НУКЛЕОТИДНЫЙ СОСТАВ РНК

Аналогом правил Чаргаффа для РНК были бы условия $\chi_A = \chi_U$ и $\chi_G = \chi_C$, однако ни один из основных типов клеточной РНК — транспортная РНК (тРНК), матричная РНК (мРНК), рибосомная РНК (рРНК) или гетерогенная ядерная РНК (гяРНК) — не подчиняется данному правилу. Это естественно, так как все они являются одноцепочечными ковалентными структурами. Правила Чаргаффа выполняются лишь для небольшого числа РНК вирусной природы, и некоторые из них оказались совершенными двухцепочечными молекулами. Нуклеотидный состав большинства РНК таков, что закручивание цепи на себя не может приводить к спариванию всех нуклеотидов. Более того, некоторые необычные основания столь сильно модифицированы, что нормальное спаривание оснований химически невозможно. Из опыта известно, что, хотя в большинстве РНК значительное количество оснований спарено, максимальная степень такого спаривания достигается редко.

Биосинтез РНК начинается с 5′-конца конденсацией двух молекул pppN (где N соответствует G или A). Процесс идет по направлению к 3′-концу до терминации. Таким образом, начальный транскрипт РНК должен содержать 5′-трифосфат и свободный нефосфорилированный 3′-конец, однако обнаружить 5′-трифосфат практически никогда не удавалось, потому что большинство типов РНК, присутствующих в значительных количествах in vivo, образуются в результате деградации более крупных предшественников РНК. Труднее сделать какие-либо выводы о 3′-конце. Если бы фосфат в этом случае удалось обнаружить, можно было бы с уверенностью сказать, что этот конец образован в результате процессинга, но 3′-фосфат в биологически активных системах почти никогда не обнаруживается. Многие известные нуклеазы могут расщеплять фосфодиэфирную связь, образуя свободный 3′-гидроксил, и поэтому нельзя однозначно сказать, в результате какого процесса — синтеза или деградации — возникает такой конец.

Вирусные РНК, как и мРНК прокариот, состоят в основном из четырех обычных оснований. мРНК же эукариот и их вирусов, как правило, содержат метилированные основания, а кроме того, у многих из них на 5′-конце имеется необычная последовательность

$$G^{5'}pppNpMp\ldots(pA)_npA^{3'}$$

Она содержит 5′-5′-трифосфатный мостик, который синтезируется после транскрипции мРНК и, вероятно, важен для трансляции. Кроме того, у многих из этих мРНК на 3′-конце имеется участок из 100 или более остатков pA. Эти остатки в большинстве (если не во всех) случаев также присоединяются после транскрипции.

5′- и 3′-концы рибосомных РНК не модифицированы, но содержат больший процент минорных оснований. В транспортных РНК доля необычных оснований и нуклеозидов может достигать 20% от общего числа. Было обнаружено, что у более развитых в эволю-

ционном отношении организмов содержание необычных нуклеозидов выше, чем у более примитивных. Все имеющиеся данные говорят о том, что эти модификации оснований реализуются после транскрипции предшественников тРНК и рРНК.

Нет никаких указаний на то, что нуклеотидный состав РНК или ДНК хоть как-то коррелирует с биологическими функциями этих молекул. Тем не менее некоторые из физических свойств двухцепочечных молекул, такие, как термостабильность и плавучая плотность, изменяются почти линейно в зависимости от мольной доли G + C. Эта связь лежит в основе многих аналитических и препаративных методов (гл. 22). Кроме того, вряд ли природа не воспользовалась сравнительной легкостью расхождения цепей двойной спирали в областях с высоким χ_{A+T} для того, чтобы связать такие области с какими-либо специфическими функциями. Так ли это на самом деле — пока неясно.

ДРУГИЕ КОМПОНЕНТЫ НУКЛЕИНОВЫХ КИСЛОТ

Помимо сахаров, фосфатов и оснований нуклеиновые кислоты содержат и другие компоненты. Так, некоторые РНК имеют участки, прочно связывающие двухвалентные катионы, и in vivo эти участки, вероятно, заняты Mg^{2+} и полиаминами. Поскольку нуклеиновые кислоты являются полианионами, они должны образовывать прочные комплексы с поликатионами соответствующих размеров. Поэтому такие полиамины, как спермин и спермидин, могут играть важную роль в функционировании ДНК или РНК in vivo. И действительно, есть данные о высокоспецифичной локализации этих соединений на тРНК in vitro. По-видимому, кроме указанных катионов нуклеиновые кислоты не несут никаких специально синтезированных и связанных с ними простетических групп. Следует иметь в виду, однако, что в клетке нуклеиновые кислоты редко существуют как свободные молекулы, они почти всегда прочно связаны со специфическими белками. Ниже мы рассмотрим основные свойства таких нуклеопротеидных комплексов.

Известен один случай специфического ковалентного взаимодействия между аминокислотами и нуклеиновыми кислотами, который осуществляется при аминоацилировании, или активации тРНК[1]. Каждый вид тРНК узнается специфической аминоацил-тРНК — синтетазой, которая катализирует следующую реакцию:

$$\text{ATP} + \text{тРНК}-\text{CpCpA}\begin{matrix}\diagup\text{OH}\\ \diagdown\text{OH}\end{matrix} + \text{NH}_3^+-\text{CHR}-\text{COO}^- \rightleftarrows$$

$$\rightleftarrows \text{AMP} + \text{PP}_i + \text{тРНК}-\text{CpCp A}\begin{matrix}\text{OH}\\ |\\ -\text{O}-\text{CO}-\text{CHR}-\text{NH}_2\end{matrix}$$

Известно, что ацилирование в некоторых случаях происходит специфически или по 2′-, или по 3′-гидроксилу концевого аденозина (две OH-группы на схеме). Однако аминокислотный остаток довольно быстро мигрирует между 2′- и 3′-положениями сахара, и, возможно, этот процесс играет какую-то роль в механизме синтеза белка. Интересно, что некоторые РНК вирусов растений содержат ковалентные структуры, сходные с 3′-концевым участком тРНК, и некоторые синтетазы могут аминоацилировать эти вирусные РНК. Значение этого явления in vivo неясно.

---

[1] В последнее время обнаружены многочисленные ковалентные соединения белков с вирусными РНК и ДНК, а также с плазмидами. Кроме того, ферменты топоизомеразы, изменяющие число сверхвитков в кольцевых ДНК, также связываются с ними ковалентно в процессе функционирования. Обычно белок присоединен к 5′-концу полинуклеотидной цепи фосфодиэфирной связью через остаток тирозина (реже — серина). — *Прим. ред.*

3.3. Первичная структура

СТРОЕНИЕ КОВАЛЕНТНОЙ ЦЕПИ

Число полинуклеотидных цепей в молекуле нуклеиновой кислоты можно определить, создав условия, при которых полностью разрушаются пары оснований (такие условия имеют место, например, в органических растворителях,при высокой температуре или в водных растворах при щелочных pH[1]), и затем измерив молекулярную массу и количество каждой из цепей. Некоторые типы структур, образуемые ковалентными цепями и обнаруженные в природных нуклеиновых кислотах или полученные искусственно, изображены на рис. 3.4.

Почти все РНК состоят из одной ковалентной цепи. Одно из возможных исключений составляет 28S-рРНК эукариот, которая, по-видимому, содержит короткую 5,8S-цепь, присоединенную нековалентно. Другим исключением является РНК некоторых опухолеродных вирусов, вероятно, существующая в виде димера из двух одинаковых цепей. Имеющиеся по этому поводу данные мы рассмотрим ниже.

Подавляющее большинство ДНК состоит из двух одиночных цепей одинаковой длины с комплементарными последовательностями, что приводит к спариванию оснований по всей длине молекулы. Иногда различие в содержании нуклеотидов в этих цепях позволяет разделять их и исследовать по отдельности. Разделение цепей in vitro показывает, что некоторые ДНК имеют одноцепочечные разрывы. Наиболее изучена в этом отношении ДНК фага T5, в которой одна цепь является целой, а другая, комплементарная ей, состоит из четырех фрагментов, точно подогнанных к первой цепи. Одни природные ДНК являются кольцевыми, а другие — линейными. Иногда наблюдаются даже сцепленные кольцевые ДНК.

Среди природных ДНК не обнаружено молекул со шпильками, образующимися в результате закручивания цепи на себя и представляющими собой совершенную двойную спираль, за исключением нескольких оснований в петле. Однако такие структуры легко можно получить искусственным путем. Точность измерения содержания нуклеотидов недостаточно высока для того, чтобы отличить их от двухцепочечных молекул, однако это можно сделать, измерив, например, молекулярную массу в денатурирующих условиях.

Длина цепей нуклеиновых кислот изменяется в огромных пределах. Самые короткие природные РНК (тРНК) состоят из 75—84 звеньев, самые длинные — РНК из класса гетерогенных ядерных РНК (гяРНК) — могут содержать до $2 \cdot 10^5$ звеньев.

Самые низкомолекулярные ДНК, обнаруженные к настоящему времени, содержат только несколько тысяч звеньев, тогда как наиболее крупные достигают длины целых эукариотических хромосом и содержат до 10^8 звеньев. Известны двухцепочечные молекулы, у которых одной цепью является РНК, а другой — ДНК. Они образуются как промежуточные формы, когда обратная транскриптаза синтезирует ДНК, комплементарную РНК. Такие комплексы возникают на короткое время — по крайней мере в ограниченной области — и при синтезе РНК на ДНК. Гибриды РНК — ДНК легко получаются in vitro путем отжига двух цепей с комплементарными последовательностями. Известно несколько случаев, когда участки РНК и ДНК ковалентно сшиваются с образованием одной цепи. Это происходит при инициации синтеза ДНК обратной транскриптазой, поскольку этот фермент использует РНК-затравку для начала синтеза новой ДНК.

Не удивительно ли, что природа, как правило, не использует молекулы, в которых рибонуклеотиды и дезоксирибонуклеотиды связаны в одну цепь? Возможно, это обусловле-

---

[1] Поскольку 2′-ОН-группа находится в *цис*-положении по отношению к фосфодиэфирной связи, в щелочных растворах РНК гидролизуется. Применение щелочи, таким образом, безопасно лишь для ДНК.

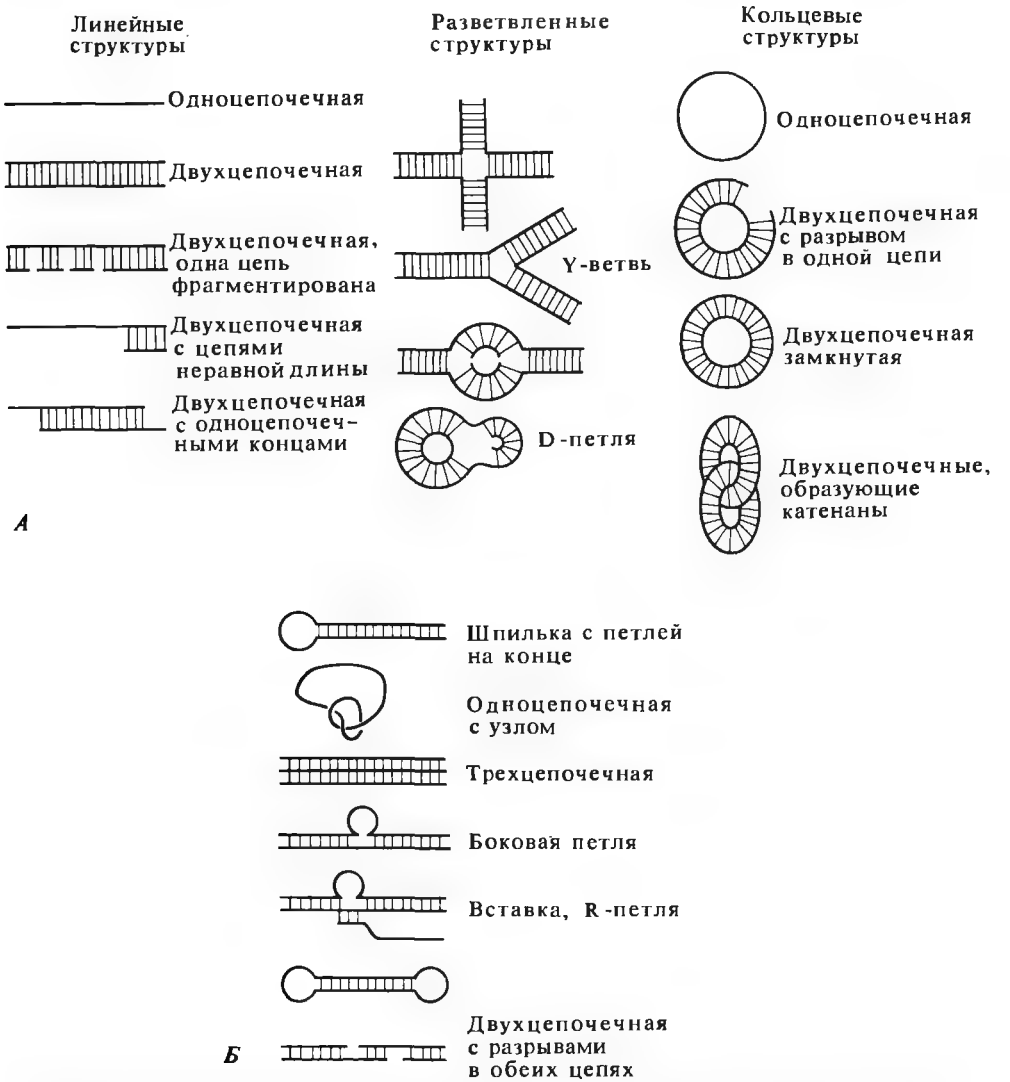


РИС.3.4. Некоторые типы структур, образуемые ковалентными цепями нуклеиновых кислот. Спиральные участки заштрихованы. *А*. Структуры, наблюдающиеся в природных нуклеиновых кислотах. *Б*. Структуры, синтезированные искусственно.

но тем, что, как будет показано ниже, рибо- и дезоксирибокомпоненты предпочитают образовывать двойные спирали с совершенно разной структурой.

НУКЛЕОТИДНЫЕ ПОСЛЕДОВАТЕЛЬНОСТИ

Последовательность оснований установлена для довольно большого числа сравнительно коротких РНК и лишь для нескольких более длинных. Все эти данные могли бы быть получены с помощью утомительного классического метода специфического нукле-

азного расщепления молекул на перекрывающиеся области, аналогично тому как были установлены аминокислотные последовательности большинства белков. Однако новые, элегантные методики значительно ускорили дело.

До самого последнего времени расшифровка нуклеотидных последовательностей ДНК представляла почти неразрешимую проблему не только из-за большого размера молекул, но также из-за отсутствия подходящих нуклеаз. Часто прибегали к транскрибированию участков ДНК с образованием РНК или же каким-либо образом встраивали определенные рибонуклеотиды в ДНК, чтобы использовать имеющиеся рибонуклеазы для частичного расщепления ДНК. Однако за последние несколько лет было открыто большое число высокоспецифичных рестриктаз. Они расщепляют двухцепочечную ДНК в определенных последовательностях, состоящих из четырех-шести оснований, и позволяют расчленить длинные молекулы на фрагменты, поддающиеся анализу. Для определения последовательностей в таких рестрикционных фрагментах были разработаны новые, очень мощные методы. Ключевым моментом в этих методах является возможность разделения (с помощью электрофореза) фрагментов ДНК, которые различаются по длине на один нуклеотид, в то время как сами фрагменты содержат до нескольких сотен звеньев; это позволяет точно локализовать места химических или ферментативных расщеплений просто измерением длины получающихся фрагментов ДНК. Эти новые методы столь эффективны, что они за короткое время дали возможность получить информацию о последовательностях ДНК, превосходящую по объему все накопленные за десятилетия данные о последовательностях белков и РНК.

Сведения о последовательности оснований в нуклеиновых кислотах позволяют глубже понять как биологические, так и физические свойства РНК и ДНК. Поскольку генетический код известен, почти всегда можно установить, какая белковая последовательность должна синтезироваться под контролем конкретной нуклеотидной последовательности. Единственная неопределенность имеется в выборе рамки считывания, или фазы триплетного кода. Здесь, однако, имеются лишь три возможности, и обычно только при правильном выборе рамки можно получить относительно протяженную последовательность между инициирующим и терминирующим кодонами. Поскольку этих кодонов несколько, а тройки букв образуют всего 64 кодона, случайная последовательность или последовательность с неправильной фазой будет кодировать скорее всего только короткие фрагменты полипептидной цепи между каждыми инициирующими и терминирующими кодонами.

Помимо собственно структурных генов в ДНК имеются области, участвующие в регуляции активности генов. Свойства этих областей до конца не выяснены, но некоторые исследованные участки обладают особой симметрией — у них имеются палиндромные последовательности в противоположных цепях. Это означает наличие точек локальной C_2- или псевдо-C_2-симметрии, как показано в примере на рис. 3.5. Другие участки, которые, возможно, являются просто промежутками между генами (спейсерами), иногда представляют собой близкие к периодическим последовательности. Хотя такие особенности позволяют идентифицировать некодирующие участки в молекулах нуклеиновых кислот, мы пока не в состоянии предсказывать функциональные свойства регуляторных участков только на основании их последовательности.

Зная нуклеотидную последовательность одноцепочечной молекулы, можно построить предположительные модели ее вторичной структуры, причем часто можно обойтись без применения специальных алгоритмов, если руководствоваться определенными правилами. Так, длинные непрерывные взаимно комплементарные участки почти всегда будут образовывать двухцепочечные шпильки. Один длинный двухцепочечный участок предпочтительнее двух более коротких с суммарной длиной, равной протяженности первого участка. Из двух эквивалентных взаимно исключающих спиральных участков более стабильным является тот, у которого выше GC-содержание. Если GC-содержание двух таких

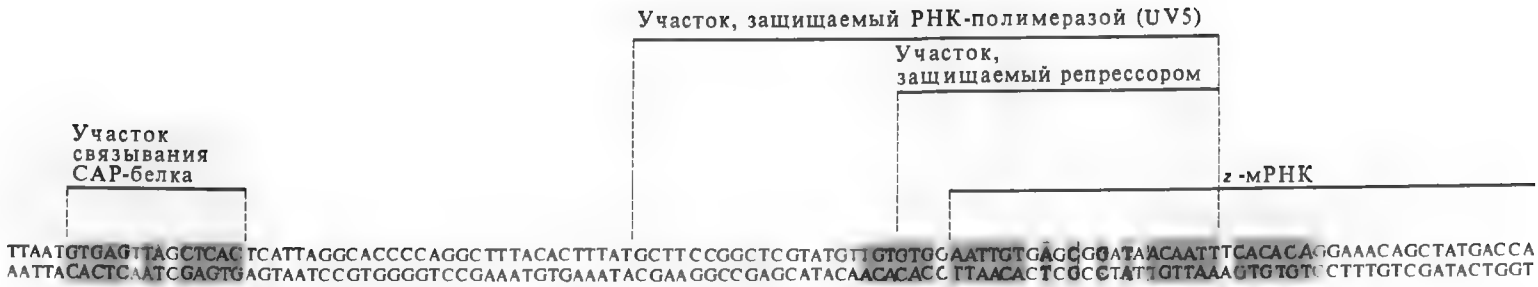


РИС.3.5. Часть нуклеотидной последовательности регуляторной области *lac*-оперона *E.coli*. Две области с псевдо-C_2-симметрией («палиндромы») закрашены. Указаны возможные места связывания РНК-полимеразы, *lac*-репрессора (ингибитора транскрипции этого оперона) и CAP-белка (общего активатора некоторых областей транскрипции). Указана часть последовательности, которая транскрибируется в начальный участок мРНК, кодирующей β-галактозидазу (z-ген). (По рисунку У. Гилберта.)

участков одинаково, более стабильным является тот, остатки которого находятся ближе друг к другу вдоль цепи (размер петли, однако, должен оставаться больше определенного минимального значения).

Эти правила могут показаться произвольными, но в гл. 23 мы детально разберем их физическую основу. О взаимодействиях нуклеотидов известно достаточно много для того, чтобы эти правила можно было использовать для количественной оценки стабильности конкретных структур. Эффективность таких простых правил особенно ярко проявляется, если обратиться к основному различию между белками и нуклеиновыми кислотами. Для белков, как мы знаем, не существует простых правил комплементарности, которые позволили бы устанавливать вторичную структуру только на основании рассмотрения парных взаимодействий между остатками.

ПРЕДПОЛОЖЕНИЯ О ВТОРИЧНОЙ СТРУКТУРЕ, ОСНОВАННЫЕ НА СРАВНЕНИИ ПОСЛЕДОВАТЕЛЬНОСТЕЙ тРНК

После установления нуклеотидной последовательности нескольких РНК разного происхождения со сходными структурами и функциями задача построения возможной модели вторичной структуры существенно упростилась. Например, для первой из молекул тРНК, последовательность которой удалось определить (Р. Холли), могли быть предложены три способа укладки цепи, приблизительно эквивалентные по отношеиию к сформулированным выше правилам. Вторая из тРНК с расшифрованной последовательностью (Г. Цахау) могла быть уложена уже только одним из трех указанных способов. Это была модель «клеверный лист», которая теперь известна как хорошая аппроксимация вторичной структуры тРНК. На сходство структуры всех тРНК указывают несколько фактов. Во-первых, разные индивидуальные тРНК и их смеси удивительно близки по своим физическим свойствам. Во-вторых, все тРНК выполняют в основном одинаковые функцин и связываются с одними и теми же участками, в частности с соответствующими участками рибосом, факторов синтеза белка и ферментов, которые могут модифицировать тРНК. На основании этих соображений модель клеверного листа была выбрана задолго до того, как удалось получить прямые физические или химические доказательства ее справедливости.

По мере расшифровки последовательности все большего числа тРНК становилось ясно, что на самом деле имеется несколько семейств «клеверных листов», похожих одно на другое и различающихся лишь в деталях. Рис. 3.6 иллюстрирует три таких семейства, охватывающих все известные тРНК, которые выполняют одинаковые функции на рибосоме. Здесь же приведены названия различных участков тРНК, которымн мы будем пользоваться в этой книге. Некоторые особенности являются одинаковыми для всех семейств; сюда относятся длина акцепторного и антикодонового стеблей, размер антикодоновой петли, часть последовательности псевдоуридиловой петли. Другие свойства, такие, как размер стебля дигидроуридиловой шпильки, строение дополиительной петли и детали дигидроуридиловой петли, сохраняются только в пределах данного семейства.

Несколько тРНК со специальными функциями не могут быть отнесены ни к одному из основных семейств. Три из них приведены на рис. 3.7. Молекула тРНК, ответственная за инициацию белкового синтеза в *E. coli*, имеет измененную длину акцепторного стебля. Эта молекула не способна выполнять многие функции, свойственные обычным тРНК. Л. Шульман в результате элегаитных экспериментов по химической модификации показал, что все эти функции восстанавливаются, если восстановить обычную длину акцепторного стебля. Эукариотическая инициирующая тРНК имеет нормальный акцепторный стебель, но у нее нет последовательности GTΨC, присутствующей у всех обычных тРНК.

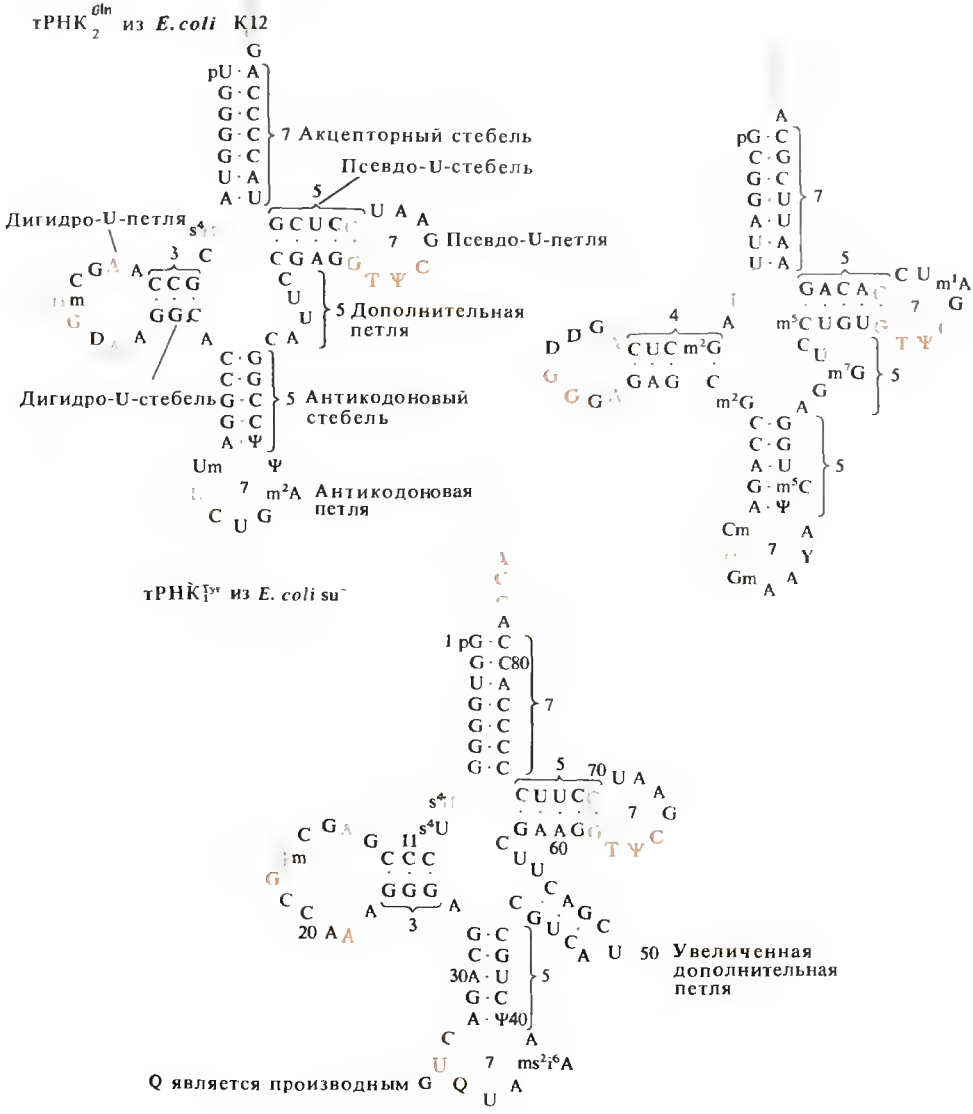


РИС.3.6. Нуклеотидные последовательности и вторичные структуры в форме клеверного листа для трех тРНК, каждая из которых представляет один из основных известных классов тРНК. Рядом с тРНК$_2^{Gln}$ из *E.coli* K12 указаны названия соответствующих участков тРНК. Нуклеотидные остатки и структурные особенности, характерные для *всех* обычных тРНК, выделены цветом. Указаны также общие для каждого класса тРНК структурные особенности.

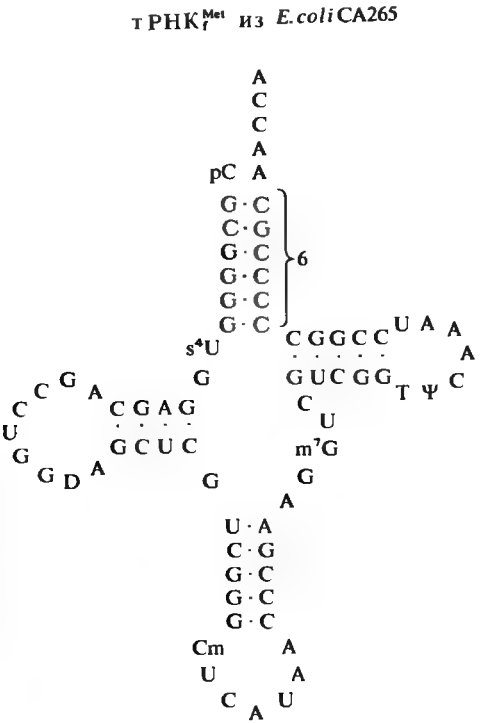


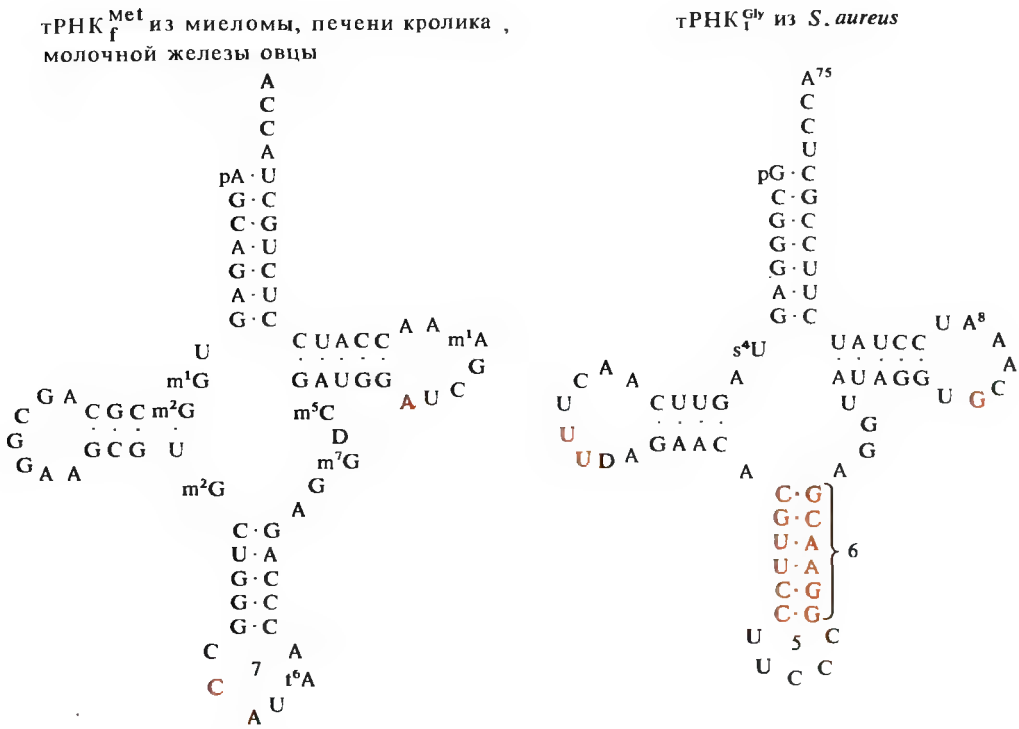
РИС.3.7. Нуклеотидные последовательности и вторичные структуры в форме клеверного листа для трех необычных тРНК. Нуклеотидные остатки и структурные особенности, которые «нарушают» правила, отмеченные на рис. 3.6, выделены цветом. Два вида тРНК$_f^{Met}$ — это инициирующие тРНК; возможно, они взаимодействуют с рибосомами иначе, чем это делают нормальные тРНК. тРНК из *S.aureus* участвует в биосинтезе пептидогликановой клеточной оболочки, и ей вообще не нужно связываться с рибосомами.

На рис. 3.7 изображена также тРНК, которая участвует в синтезе пептидогликановых структур клеточной оболочки, а не в синтезе белков на рибосомах. У этой молекулы отсутствуют многие черты, типичные для других тРНК.

ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ И АНАЛИЗ СТРУКТУРЫ И ФУНКЦИИ

Знание нуклеотидных последовательностей создает необходимые предпосылки для использования многих экспериментальных методов. Эти данные нужны для успешного применения метода химической модификации с целью исследования классов доступных и недоступных для реакции звеньев (см. обсуждение третичной структуры тРНК ниже в этой главе). Информация о последовательности необходима для анализа данных, получаемых с помощью различных химических или физических зондов, которые позволяют оценивать расстояние между определенными звеньями, участвующими, например, в образовании внутримолекулярных сшивок. Она может использоваться также в сочетании со спектроскопическими или другими физическими данными для того, чтобы попытаться установить, какие из предполагаемых спиральных участков присутствуют в нативной молекуле. Эти методы рассмотрены подробно в гл. 24.

Знание нуклеотидных последовательностей оказывается крайне важным и для установления характера взаимодействия различных нуклеиновых кислот друг с другом. Например, данные о последовательности в антикодоновой петле тРНК позволяют предсказать, какой кодон мРНК стимулирует связывание тРНК с рибосомой путем спаривания с триплетом оснований антикодоновой петли тРНК. Здесь, однако, строгие правила спари-



вания оснований несколько смягчаются. 5′- и центральной нуклеотиды кодирующего триплета (соответствующие 3′- и центральному нуклеотидам антикодона) образуют обычные пары оснований (исключение составляют инициирующие тРНК), а для третьего нуклеотида возможен более широкий набор взаимодействий:

Антикодон тРНК, 3′-основание	G A C U I
Кодон мРНК, 5′-основание	U, C U G A, G A, U, C

Способность инозина (I) образовывать пары с тремя разными комплементарными к нему остатками подтверждена физическими исследованиями на простых модельных системах. Энергия взаимодействия G · U остается пока несколько неопределенной. Ясно лишь, что GU-пара, окруженная в двойной спирали правильными парами, ослабляет двойную спираль в гораздо меньшей степени, чем другие неправильные пары (такие, как A · C или G · A). Тем не менее неясно — по крайней мере для модельных комплексов в растворе — повышает или понижает энергию спирали такая пара.

Рассмотрим еще один пример, когда особенности взаимодействия между нуклеиновыми кислотами были предсказаны на основании анализа их последовательностей и лишь позднее получили экспериментальное подтверждение. На 3′-конце 16S-рРНК из *E. coli* имеется участок, до некоторой степени комплементарный последовательностям, расположенным на 5′-конце участка инициации белкового синтеза во многих вирусных РНК, которые функционируют как мРНК (рис. 3.8). Взаимодействие этих участков используется, очевидно, для узнавания области инициации и ориентации мРНК в правильном стартовом положении. Разные участки инициации обладают разной степенью комплементарности к 16S-рРНК, и вполне возможно, что они узнаются с разной вероятностью. Следо-

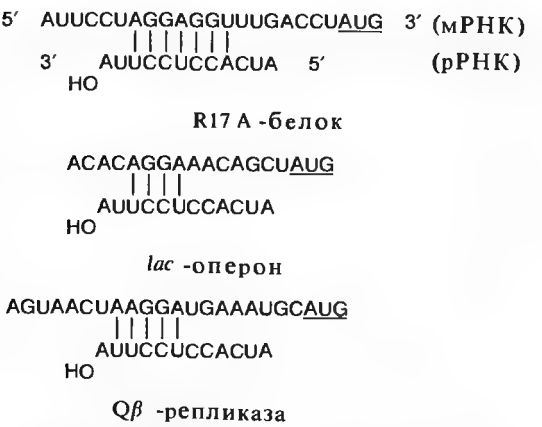


РИС.3.8. Комплементарные последовательности, расположенные у 3′-конца 16S-рРНК из *E.coli* в направлении 5′-конца от AUG-участка, с которого начинается синтез белка во многих фаговых РНК и мРНК *E.coli*. (Из работы Joan Steitz и Karen Jakes.)

вательно, это взаимодействие может играть роль в узнавании и влиять на частоту использования конкретных мРНК. Эти интересные предположения, однако, еще нуждаются в подтверждении.

МЕТОДЫ ПРЯМОГО СРАВНЕНИЯ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

Сравнение последовательностей является мощным методом исследования и часто оказывается весьма полезным, поскольку его удается проводить даже в тех случаях, когда детальные последовательности неизвестны. Определение степени гибридизации, которая представляет собой меру способности одиночных цепей объединяться в двухцепочечные структуры, может использоваться для предварительного анализа гомологии последовательностей. Для выявления стабильных двухцепочечных областей, образованных одноцепочечными ДНК или РНК из разных источников, используются различные методы. Особенно информативна кинетика такой рекомбинации (гл. 23).

Исследования сходства последовательностей в гибридных ДНК или гомологии между различными природными ДНК могут быть выполнены — с не очень высоким разрешением — с помощью электронной микроскопии. Хотя современные методы ограничивают разрешение 50 — 100 парами оснований, многие интересные вопросы удается выяснить и на этом уровне. Этот подход значительно проще, чем химическое определение последовательностей из тысяч нуклеотидов.

Предположим, что последовательности двух ДНК идентичны по всей длине молекулы, за исключением одного непрерывного участка длиной в 100 или более пар. Если эти две ДНК смешать, подвергнуть денатурации, а затем ренатурации (рис. 3.9), мы получим некоторое количество гибридных структур, называемых гетеродуплексами. Они менее стабильны, чем нормальные гомодуплексы, поскольку спаривание оснований в них несовершенно. В принципе гетеродуплексы должны постепенно диссоциировать, а разделенные нити рекомбинировать с образованием гомодуплексов. Однако при температурах, при которых двойная спираль является стабильной, этот процесс сопряжен с такой гигантской энергией активации, что значительные изменения произойдут только в космологическом масштабе времени. Участки с неспаренными основаниями, таким образом, бу-

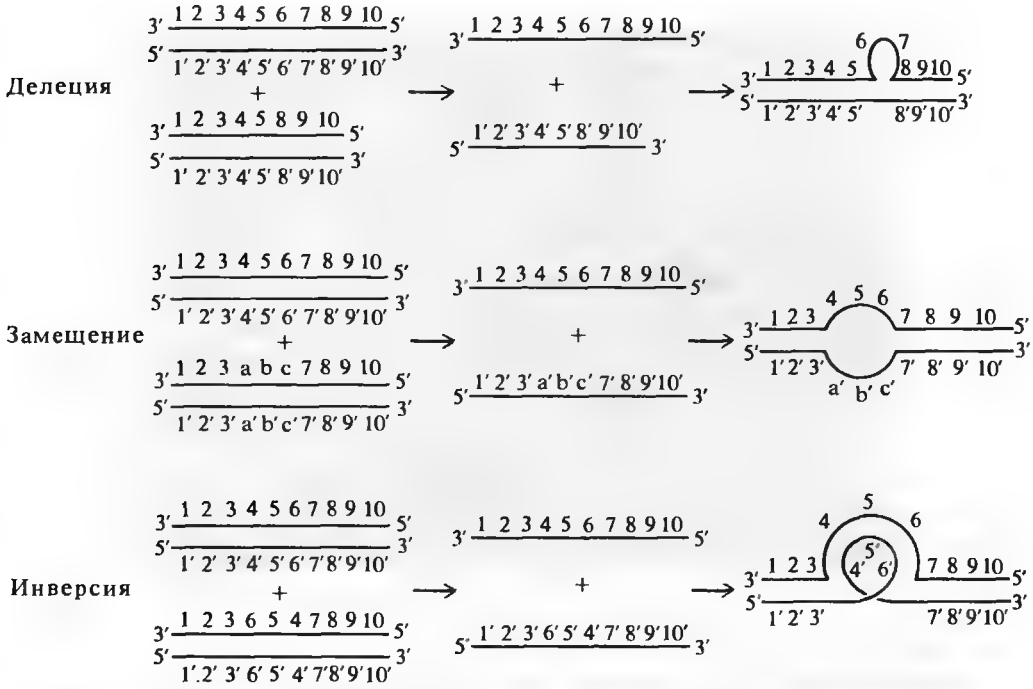


РИС.3.9. Схематическое изображение плавления ДНК и образования гетеродуплекса при ренатурации. Показана только одна из двух эквивалентных пар гетеродуплексов, которые могут образовываться в каждой из рассмотренных смесей ДНК.

дут оставаться фиксированными, и их можно будет отличить от двухцепочечных участков на электронных микрофотографиях по образуемым ими петлям.

На рис. 3.9 изображены петли трех типов, которые наблюдаются в зависимости от того, обусловлено ли различие в последовательностях делециями, замещениями или инверсиями. Гибриды между двумя молекулами ДНК образуются даже в том случае, если их последовательности совпадают лишь на коротком участке. Таким образом, электронная микроскопия может быть использована, например, для выяснения того, имеют ли определенные фрагменты ДНК одинаковый ген, и для локализации этого гена.

На рис. 3.10 приведена электроиная микрофотография гибрида ДНК — РНК. Молекула была визуализирована путем обработки препарата белком гена 32 бактериофага Т4. Этот белок имеет большую константу связывания только с одноцепочечной ДНК, поэтому одноцепочечные участки кажутся утолщенными из-за того, что они покрыты белком. Это позволяет различить одиночные и двойные цепи под микроскопом. После отжига раствора РНК и ДНК под электронным микроскопом выявляются образовавшиеся гетеродуплексы, что позволяет картировать гены, кодирующие эти РНК. Следует отметить, что очень короткие РНК, такие, как тРНК, образуют слишком короткие гетеродуплексы, чтобы их можно было увидеть под электронным микроскопом. В этих случаях для визуализации места связывания с ДНК к молекулам тРНК присоединяют электронно-микроскопический маркер (например, электроноплотный белок ферритин или шарик из полистиролового латекса).

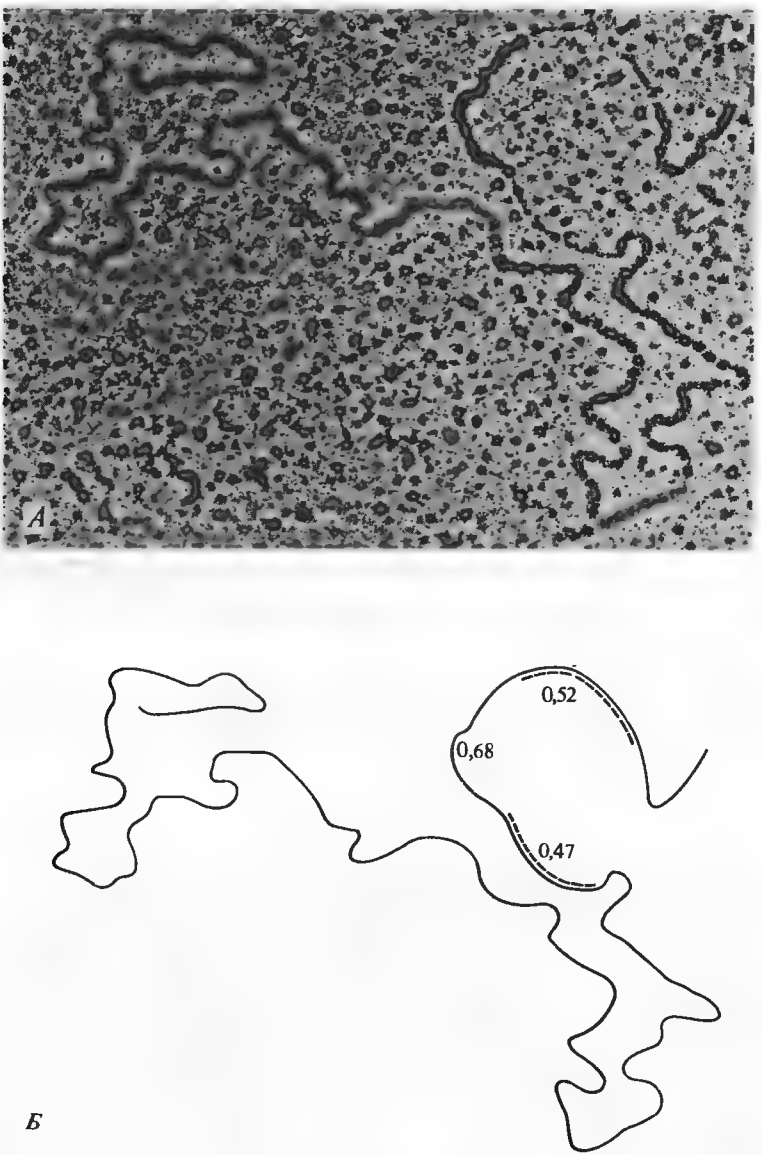


РИС.3.10. Гибрид РНК — ДНК. *А*. Электронная микрофотография. *Б*. Схематическое изображение той же молекулы. ДНК является кодирующей цепью плазмиды pSp17, содержащей два соседних гистоновых гена морского ежа. С ДНК гибридизованы молекулы мРНК этих гистонов. Двухцепочечные участки кажутся на фотографии более тонкими, потому что они не связаны с белком гена 32, который покрывает всю остальную ДНК. Числа указывают длины двух генов и участка между ними в тысячах нуклеотидных остатков. [Данные предоставлены М. Ву и Н. Девидсоном. Cell, **9**, 163 (1976).]

В большинстве случаев эффективное использование техники гетеродуплексного картирования, которую мы только что описали, требует аккуратного измерения длин одно- и двухцепочечных участков ДНК. Измерение расстояния между различными петлями или особенностями гибридов позволяет построить физическую карту ДНК. Если генетическое значение делеции или локализация отдельных генов известны, физическая карта может быть использована для построения генетической карты. Метод является особенно мощным потому, что он позволяет локализовать гены, используя лишь выделенные транскрипты РНК, даже когда не удается обнаружить мутантные фенотипы. В принципе он позволяет картировать даже области ДНК, которые вовсе не были транскрибированы. Любой изолированный или синтезированный фрагмент ДНК может быть гибридизован с образованием двухцепочечной структуры и затем идентифицирован с помощью электронного микроскопа. Длины соответствующих участков можно измерить абсолютным методом, если в поле микроскопа имеется стандарт длины для калибровки увеличения. При некоторых условиях с помощью электронного микроскопа удается получить контурную длину ДНК, находящуюся в прекрасном соответствии с теми размерами, которые вычисляют исходя из известной геометрии двойной спирали. Однако контурная длина зависит от процедуры нанесения ДНК на подложку, которая поддерживает образец в микроскопе. Поэтому на практике для более точного измерения длин к каждому образцу добавляют молекулы ДНК известной длины и используют их в качестве внутреннего стандарта.

Многие аспекты электронной микроскопии остаются, вероятно, до сих пор в такой же степени искусством, как и наукой. Чтобы расправить молекулу на подложке как можно более равномерно, нужна большая тщательность. В тех точках, где цепь ДНК пересекает сама себя, анализ ее хода становится неоднозначным. Поэтому электронные микрофотографии компактных, переплетенных молекул обычно малоинформативны. Многие методики не позволяют получить столь четкого различия между одно- и двухцепочечными областями, как на рис. 3.10.

Нужно иметь в виду, что под электронным микроскопом исследуют отдельные молекулы образца, и их конкретные структурные особенности могут оказаться просто результатом флуктуации, которая произошла в молекуле на подложке. Так, пересечения цепи легко могут имитировать петли; изгибы, различимые лишь на пределе разрешения, можно вообще пропустить, и молекула будет казаться короче, чем она есть в действительности. Поэтому необходимо просматривать много молекул и использовать статистические критерии для оценки любых наблюдаемых структурных особенностей.

В большинстве случаев возникает неопределенность, связанная с химической идентичностью двух концов линейной молекулы. Это эквивалентно неопределенности физического направления последовательности в кольцевой молекуле. Предположим, например, что картируется делеционная петля в середине цепи. В этом случае a priori неизвестно, является ли данный конец 3′- или 5′-концом цепи, которая несет участок, комплементарный пропущенному. Устранить эту неопределенность можно или с помощью генетических данных, или — более прямым путем — добавлением в среду специфических ферментов, вызывающих частичную деградацию цепи ДНК с определенного конца. Если гены картируются с помощью гибридов РНК — ДНК, иногда в распоряжении исследователя оказываются предшественники РНК, имеющие дополнительные основания или на 3′-, или на 5′-конце. В этих случаях можно определить физическое направление цепи.

3.4. Вторичная структура

Говоря о первичной структуре нуклеиновых кислот, мы не могли обойтись без использования моделей вторичной структуры как способа изображения укладки цепи или сравнения последовательностей. Обычно при расшифровке последовательностей новых одноцепочечных молекул результаты представляют в виде некоторой эстетически привлекатель-

ной схемы, согласующейся в первом приближении с приведенными выше правилами. Оставляя в стороне вопрос об адекватности такого представления молекулы, отметим, что оно весьма полезно как мнемонический прием для описания общих черт последовательности. Если же правила спаривания определены строго, количество информации, необходимое для точного определения последовательности, уменьшится приблизительно вдвое.

Рассмотрим более детально количественные аспекты вторичной структуры нуклеиновых кислот. Большинство имеющихся здесь данных получено в результате рентгеноструктурных исследований волокон двухцепочечных нуклеиновых кислот (гл. 14). Одноцепочечные компоненты, как правило, не дают хороших волокон, поэтому данные об их вторичной структуре являются менее прямыми (гл. 22).

СПАРИВАНИЕ ОСНОВАНИЙ — ГЛАВНАЯ ОСОБЕННОСТЬ ДВУХЦЕПОЧЕЧНЫХ НУКЛЕИНОВЫХ КИСЛОТ

Некоторые особенности двухцепочечных полинуклеотидов непосредственно указывают на то, что их организация радикально отличается от организации белковых молекул. В двухцепочечной структуре определенные группы оснований нуклеиновых кислот оказываются практически недоступными для различных химических агентов. Это заставляет предположить, что основания расположены внутри структуры. Напротив, все боковые группы белков локализованы снаружи образуемых молекулами спиралей или слоев (см., например, рис. 1.1). Все нуклеиновые кислоты независимо от последовательности способны образовывать одинаковую двухцепочечную структуру. В белках же последовательность является решающим фактором, от которого зависят само наличие определенной вторичной структуры и ее тип.

Идея о такой структурной организации нуклеиновых кислот, при которой все боковые группы расположены внутри структуры и которая может «приспосабливаться» к любой последовательности боковых групп, кажется парадоксальной. Ключ к пониманию этого обстоятельства дали правила Чаргаффа для содержания различных оснований, но нужно признать, что они относились только к молекуле нуклеиновой кислоты в целом, а не к отдельным цепям. Пытаясь объединить всю эту информацию, Джеймс Уотсон и Френсис Крик выдвинули идею о специфическом взаимодействии между комплемеитарными основаниями: A с T или U, G с C. Напомиим, что в то время не было прямых данных о специфическом взаимодействии между основаниями. Кроме того, выбор именно этих, а не других возможных пар оснований был обусловлен тем, что Уотсон и Крик остановились на правильных кето-амино-таутомерных формах оснований, обнаруженных лишь незадолго до этого и не являющихся еще общепринятыми.

Важиейшей особенностью AT- и GC-пар, на которую обратили внимание Уотсон и Крик, была не просто хорошая геометрия образуемых водородиых связей. Другие конфигурации спаривания основаиий также обладают этим свойством и действительно превалируют в комплексах между изолированными мономерами (гл. 6). Особенностью предложенной Уотсоном и Криком схемы является то, что расстояния между местами присоединения оснований к сахарам оказываются практически одинаковыми для AT- и GC-пар. Поэтому одна и та же геометрия двойной спирали одииаково хорошо подходит для обоих типов пар без какого-либо иарушения или потери симметрии (рис. 3.11). Второй важной особенностью модели Уотсона — Крика является существование осей псевдосимметрии C_2, лежащих в плоскости пар оснований. Поворот вокруг этих осей переставляет $N^9—C^{1'}$-гликозидную связь пурина с $N^1—C^{1'}$-связью комплементарного пиримидина. Эти оси существуют как в AT-, так и в GC-парах (рис. 3.11), откуда следует, что геометрия AT-, TA-,

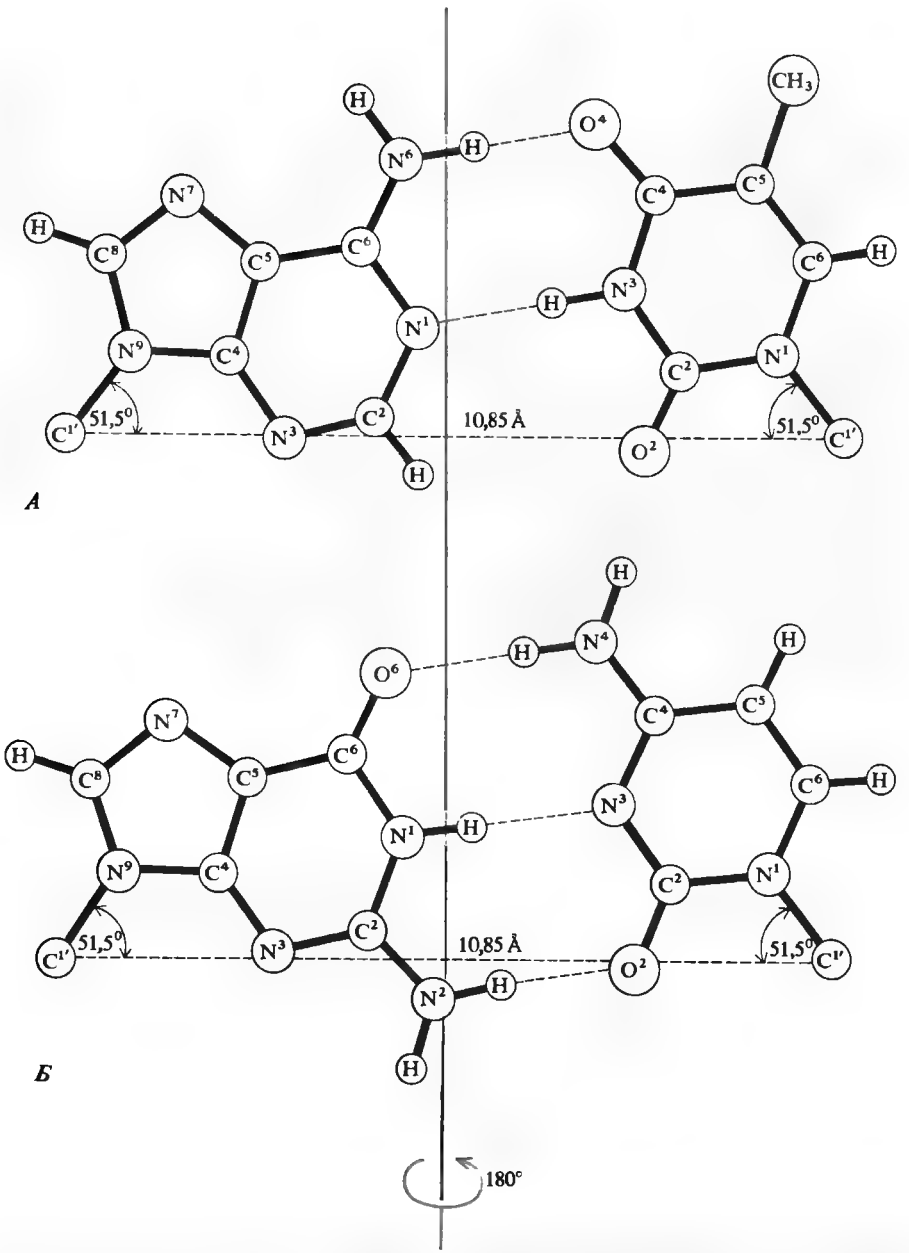


РИС.3.11. Уотсон-криковские пары оснований AT и GC. Важной особенностью этих структур является постоянство расстояний $C^{1'} — C^{1'}$ и углов, которые связи $N^9 — C^{1'}$ и $N^1 — C^{1'}$ образуют с вектором $C^{1'} — C^{1'}$. Указана ось псевдосимметрии C_2. [S. Arnott, S.D. Dover, A.J. Wonacott, Acta Cryst., **B25**, 2192 (1969)].

Таблица 3.1

СТРУКТУРА ПОЛИНУКЛЕОТИДНЫХ ДВОЙНЫХ СПИРАЛЕЙ

Полинуклеотид	Относительная влажность, %	Число остатков на виток	Расстояние между остатками вдоль оси спирали, Å	Угол между плоскостью оснований и осью спирали, град	Двугранный угол между плоскостями оснований, град	Конформация сахара
ДНК-А, Na^+	75	11	2,55	70	16	$C^{3'}$-*эндо*
ДНК-В, Na^+	92	10	3,46	-	-	-
ДНК-В, Li^+	66	10	3,37	88	5	$C^{2'}$-*эндо*
ДНК-С, Li^+	66	9,3	3,32	84	10	$C^{2'}$-*эндо*
Гибрид ДНК—РНК, Na^+	75	11	2,62	70	-	-
Фрагменты дрожжевой РНК	75	10 или 11	2,9 или 2,64	-	-	-
Реовирус, α- или β-формы	75	10 или 11	3,0 или 2,73	-	-	-
Poly (rA · rU)		10 или 11	3,1	-	-	-
Poly (rI · rC)		12	3,0	-	-	-
Poly (rCH^+ · rC)[1]		12	3,11	-	-	-
Poly (rAH^+ · rAH^+)[1]		8	3,8	-	-	-

[1] Структуры с параллельными цепями.

GC- и CG-пар одинакова. Каждое из четырех звеньев может занимать любое положение в двойной спирали. Основания в каждой цепи двойной спирали имеют в точности одинаковое окружение.

СТРУКТУРА ДВОЙНЫХ СПИРАЛЕЙ НУКЛЕИНОВЫХ КИСЛОТ

Фиксированная структура уотсон-криковских пар оснований допускает большое многообразие форм двойной спирали. В табл. 3.1 суммированы параметры ряда форм, которые были обнаружены экспериментально. На рис. 3.12 и 3.13 схематически изображена молекулярная структура А- и В-форм ДНК, а на рис. 3.14 показаны объемные модели ДНК-В и РНК-11. Эти рисунки следует изучить более внимательно параллельно с текстом этого раздела.

Приведенные структуры имеют некоторые общие особенности. Все эти структуры являются правыми спиралями. Все имеют оси симметрии 2-го порядка, которые представляют собой истинные оси симметрии для фосфодиэфирного остова и оси псевдосимметрии для самих пар. Поскольку они являются осями вращения, фосфодиэфирные остовы должны быть ориентированы в противоположных направлениях. Концы двойной спирали химически и структурно идентичны, так как каждый содержит по одному 3′- и 5′-

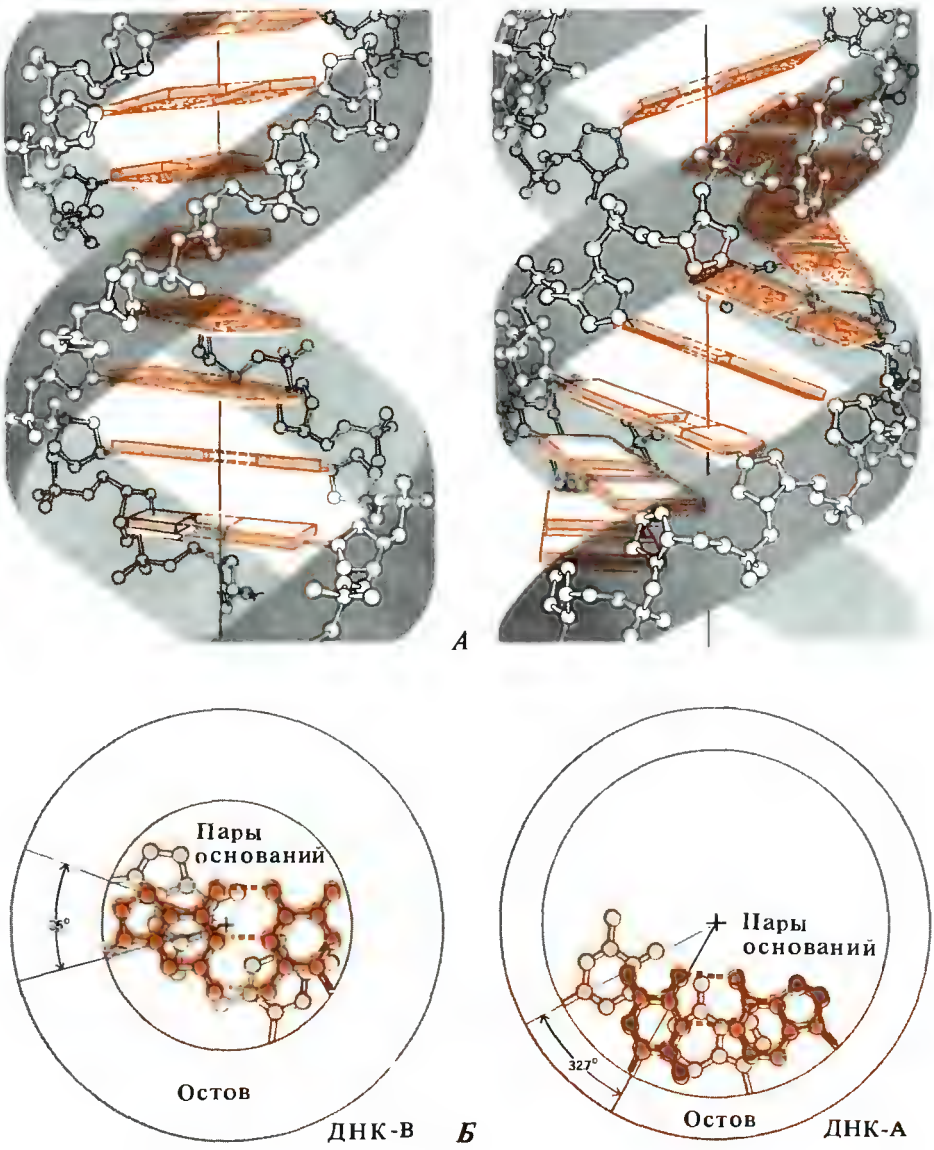


РИС.3.12. Двойная спираль ДНК в В- и А-формах. *А.* Оси спиралей лежат в плоскости рисунка. На этой схематической иллюстрации освования изображены в виде брусков, а атомы остова, за исключением атомов водорода, изображены каждый по отдельности в виде кружков. Отчетливо видно, что плоскости пар оснований в А-ДНК наклонены к ее оси. *Б.* Оси спиралей (+) перпендикулярны плоскости рисунка. В ДНК-В ось проходит прямо между основаниями, образующими пару, тогда как в ДНК-А пары оснований, поворачиваясь вокруг оси (если перемещаться вдоль последней), как бы образуют полый цилиндр. (Рисунки Ирвинга Гейса. Выполнены при участии В.Сенжера.)

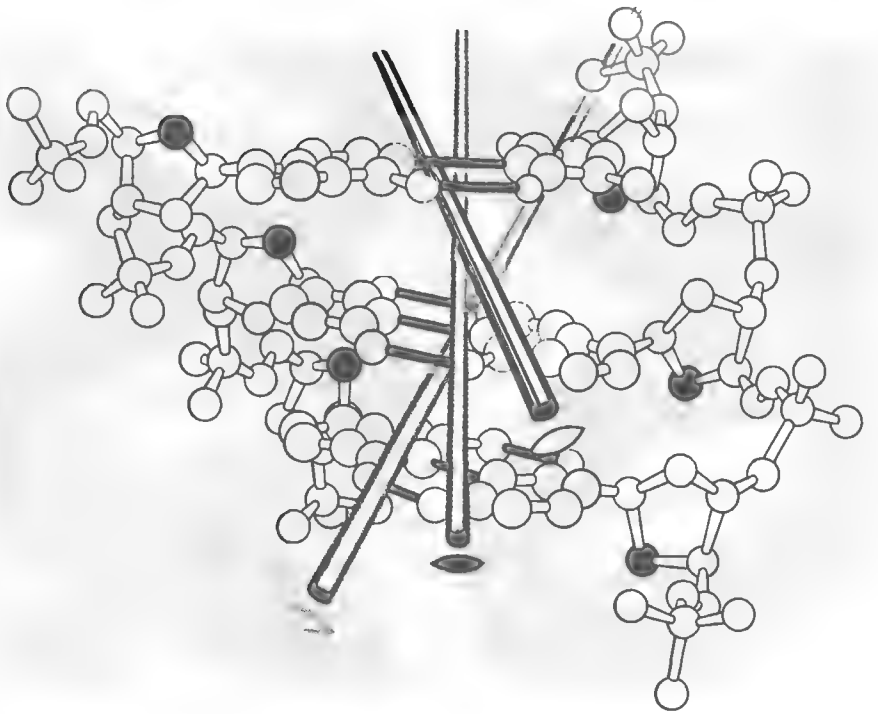


РИС.3.13. Оси симметрии в ДНК-В. Показаны две оси симметрии 2-го порядка (○) и ось псевдосимметрии C_2 (●). (Рисунок Иргвинга Гейса.)

концу одиночных цепей. В Дополнении 3.1 представлено простое химическое доказательство того, что полинуклеотидные цепи со спаренными основаниями должны быть антипараллельными.

Дополнение 3.1

КАК ДОКАЗАТЬ, ЧТО ЦЕПИ В ДВУХЦЕПОЧЕЧНОМ КОМПЛЕКСЕ АНТИПАРАЛЛЕЛЬНЫ

Одной из важнейших особенностей всех описанных до сих пор двойных спиралей является антипараллельность образующих их цепей. То, что такая структура действительно реализуется для молекул ДНК и РНК в растворе, можно показать с использованием простой и элегантной методики анализа последовательностей, разработанной Артуром Корнбергом. Молекулы ДНК и РНК синтезируются ферментативным путем из 5′-трифосфатов. Ниже приведена схема, показывающая, как синтез из α-^{32}P-меченых предшественников можно использовать для анализа частоты встречаемости различных комбинаций соседних вдоль цепи звеньев. Для определения встречаемости соседних с A звеньев эта схема выглядит следующим образом:

$$ppp^*A + pppT + pppG + pppC \xrightarrow{\text{ДНК-полимераза I}} pGp^*ApCpTp^*Ap^*ApCpCp^*A\ldots$$

Антипараллельность цепей имеет решающее значение при спаривании оснований в одноцепочечных молекулах (таких, как тРНК). Параллельное спаривание не может обеспечить образование шпилек. Двойные спирали, изображенные на рис. 3.12, 3.13 и 3.14, имеют второй набор осей псевдосимметрии C_2, проходящих между соседними парами связанных водородными связями оснований. Для самокомплементарных последовательностей, таких, как GpC или ApT (или палиндромных последовательностей, аналогичных приведенным на рис. 3.5), оси, проходящие между двумя центральными парами, являются истинными осями симметрии, поскольку поворот на 180° не изменяет положения остова и переводит пары оснований в идентичные положения.

Еще одной особенностью всех двойных спиралей является стэкинг (межплоскостное взаимодействие) оснований. Хотя относительное расположение соседних пар оснований меняется, во всех случаях расстояние между плоскостями соседних пар составляет около 3,5 Å. Это соответствует вандерваальсовым радиусам плоских ароматических соединений и является наименьшим расстоянием, на котором еще не происходит взаимного отталкивания химически не связанных групп.

Несмотря на все отмеченное сходство, структура двухцепочечных молекул нуклеиновых кислот во многом различается. Эти различия касаются числа пар оснований на виток двойной спирали, угла между парами оснований и осью спирали, угла между плоскостями, в которых лежат основания, диаметра спирали, характера желобков спирали. Некоторые из этих особенностей видны на проволочных моделях, а желобки лучше всего выявляются на объемных моделях, показанных на рис. 3.14.

Рассмотрим наиболее распространенные структуры. ДНК в B-форме обнаружена в волокнах, содержащих ионы Na^+, при высокой относительной влажности. Это дает основания полагать, что она соответствует форме, присутствующей в обычных водных растворах, и теперь известно, что это действительно так (некоторые доказательства приведены в гл. 22). ДНК в B-форме содержит 10 пар оснований на виток двойной спирали. Пары оснований являются практически плоскими и перпендикулярны оси спирали. Двойная спираль имеет два желобка: малый желобок, лежащий между 1′-углеродными атомами сахаров каждой из цепей, и большой желобок, расположенный напротив малого с другой стороны спирали. Диаметр спирали равен примерно 22 Å.

B-форма обычно считается уотсон-криковской структурой. Тем не менее первоначальная модель Уотсона и Крика, хотя и была основана на данных, полученных для B-формы

$$\ldots \xrightarrow{\text{ДНКаза}} Gp^* + Ap + Cp + Tp^* + Ap^* + Ap + Cp + Cp^* + \ldots$$

Долю звеньев каждого типа, предшествующих A, можно вычислить по данным об относительном распределении метки между четырьмя нуклеотидами. Исходя из этого и зная содержание оснований, проделав аналогичные эксперименты для каждого из четырех типов меченых трифосфатов, можно определить частоты встречаемости всех 16 комбинаций ближайших соседей. Для антипараллельного уотсон-криковского спаривания оснований должны выполняться следующие соотношения для мольных долей χ различных соседей:

$$\chi_{CpT} = \chi_{ApG}, \qquad \chi_{ApC} = \chi_{GpT}$$

Для параллельных цепей эти ограничения в общем случае не выполняются. Вместо них должны иметь место соотношения

$$\chi_{CpT} = \chi_{GpA}, \qquad \chi_{ApT} = \chi_{TpA}$$

Используя этот принцип, было легко показать, что цепи двухцепочечных ДНК, РНК и гибридов антипараллельны.

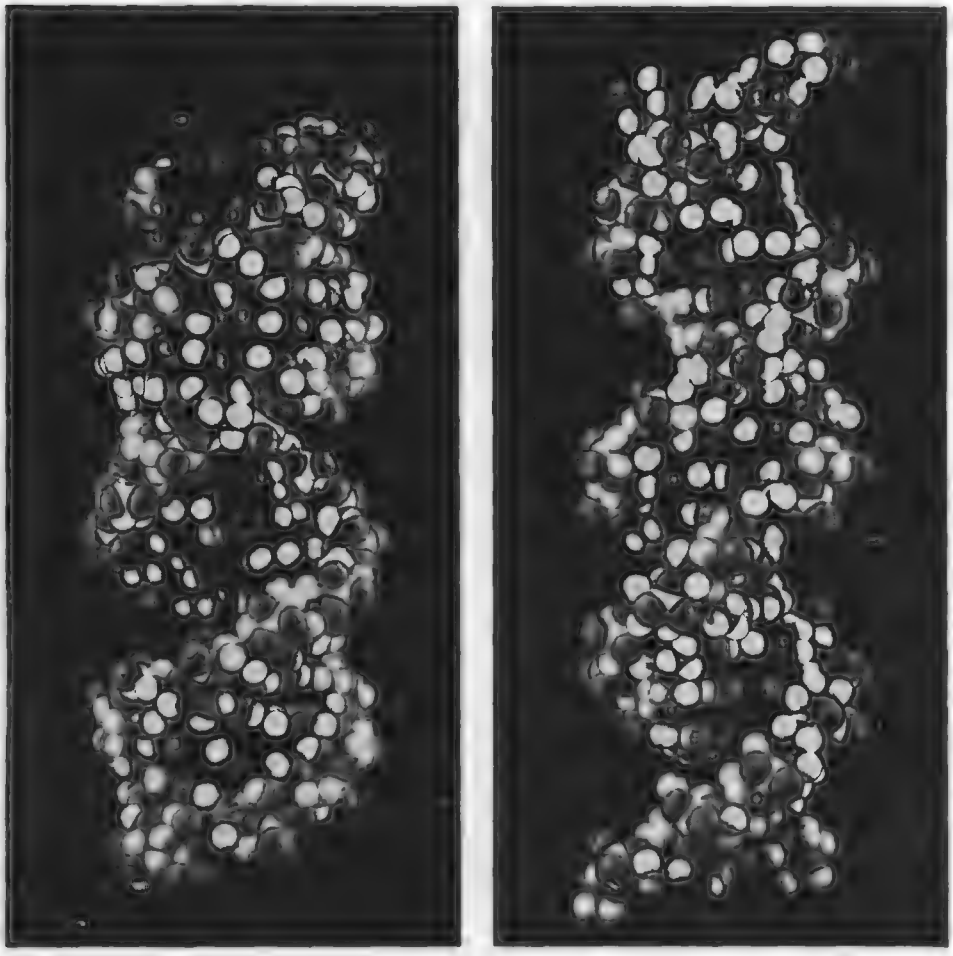


РИС.3.14. Объемные модели двойных спиралей ДНК-В (*справа*) и РНК-11 (*слева*). Хорошо видна существенная разница между этими двумя формами в шаге спиралей и в размере желобков. (Фотографии предоставлены Сунг Хоу Кимом.)

ДНК, отличалась от принятой теперь наличием лишь двух водородных связей между G и C, а также тем, что по многим своим структурным деталям была ближе к A-, чем к B-форме. Для построения принятой в настоящее время геометрии B-формы потребовались более точные экспериментальные данные и развитие методов анализа рентгенограмм.

A-форма ДНК наблюдается в волокнах в присутствии ионов Na^+ при относительной влажности ниже 75%. Она имеет 11 пар на виток и диаметр на несколько ангстрем больший, чем B-форма. Пары оснований тоже почти плоские, но в этой форме они наклонены на 20° к плоскости, перпендикулярной оси спирали. Таким образом, B-форма похожа на обычную винтовую лестницу, а A-форма — на винтовую лестницу, ступеньки которой наклонены к центру. A- и B-формы переходят одна в другую при изменении относительной влажности. Если процесс происходит быстро и без разрушения упорядоченной упа-

ковки волокон, переход может осуществляться без расхождения цепей двойной спирали. Это является веским доводом в пользу того, что А- и В-формы имеют один и тот же знак спирали. В растворе В-форма, существующая в обычных условиях, может быть переведена а А-форму уменьшением активности растворителя, например добавлением этанола в водный раствор.

Находятся ли природные ДНК в А-форме — неизвестно, но можно с достаточной определенностью сказать, что гибриды РНК — ДНК существуют в А-, а не в В-форме даже в обычных водных буферных растворах. С помощью рентгеноструктурного анализа было исследовано множество двойных спиралей РНК. Оказалось, что структурой, наилучшим образом согласующейся с экспериментальными данными, является спираль РНК-11. Она имеем 11 пар оснований на виток, а угол наклона плоскости пары оснований к плоскости, перпендикулярной оси спирали, составляет 13 — 14°. Тем не менее дифракционные данные, полученные при исследовании таких РНК, недостаточно полны и точны, чтобы исключить альтернативную спираль — РНК-10. Последняя имеет 10 пар оснований на виток, отклоняющихся на 10° от перпендикуляра к оси спирали. Структура РНК-10 занимает как бы промежуточное положение между А- и В-формами ДНК.

Рассмотрим вкратце еще одну форму двойной спирали полинуклеотидов — С-форму ДНК. Здесь число пар оснований на виток равно 9,3, плоскость оснований наклонена под небольшим — около 6° — углом к плоскости, перпендикулярной оси спирали, а плоскости самих оснований также несколько наклонены друг относительно друга. С-форма, обнаруженная в литиевой соли ДНК при низкой влажности, — это в действительности разновидность В-формы. Существуют и другие варианты как А-, так и В-форм, так что они представляют скорее семейства структур, чем строго определенные конформации двойной спирали.

СТРУКТУРНЫЕ РАЗЛИЧИЯ МЕЖДУ РНК И ДНК

РНК никогда не наблюдалась в В-форме, и это вполне понятно. Присутствие 2′-гидроксила изменяет конформацию рибозного кольца таким образом, что переход в В-форму становится для РНК невозможным. Из рентгеноструктурных исследований различных мононуклеотидов известно, что четыре из пяти атомов пиранозного кольца обычно лежат в одной плоскости, в то время как пятый ($C^{2'}$ или $C^{3'}$) выходит из плоскости. Это приводит к четырем различным конформациям сахарного кольца, показанным на рис. 3.15. В структуре $C^{2'}$-*эндо* 2′-углеродный атом лежит над плоскостью пиранозного кольца, т.е. с той же стороны от плоскости, что и основание и 5′-углерод. Очень похожа на нее структура $C^{3'}$-*экзо*. Напротив, в структуре $C^{3'}$-*эндо* 3′-углеродный атом лежит выше плоскости пиранозного кольца, а $C^{2'}$-*экзо*-конформация очень похожа на нее.

Преимущественные конформации рибозы, полученные из рентгеноструктурых исследований, оказываются различными для разных полинуклеотидных спиралей и мононуклеотидов. В спиралях РНК рибоза находится почти исключительно в $C^{3'}$-*эндо*-конформации. Для мономерных рибонуклеотидов конформации $C^{3'}$-*эндо* и $C^{2'}$-*эндо* встречаются примерно в одинаковом числе структур, исследованных с помощью рентгеноструктурного анализа кристаллов или методом ЯМР в растворе. Дезоксирибоза в спиралях ДНК может находиться как в конформации $C^{2'}$-*эндо*, так и $C^{3'}$-*эндо*, но в мономерных дезоксирибонуклеотидах почти всегда обнаруживается $C^{2'}$-*эндо*-конформация. Пираноза в составе В-формы должна находиться в $C^{2'}$-*эндо*-конформации. Полинуклеотиды в А-форме и РНК-10 или 11 содержат сахара в $C^{3'}$-*эндо*-конформации. Точная величина проигрыша в энергии при переходе от $C^{3'}$-*эндо* к $C^{2'}$-*эндо*-конформации для рибонуклеотидов неизвестна, однако перевод РНК в В-форму связан с огромными энергетическими затратами, поскольку при этом необходимо перевести в состояние с более высокой энергией все звенья цепи. По-видимому, РНК не может находиться в В-форме из-за энергети-

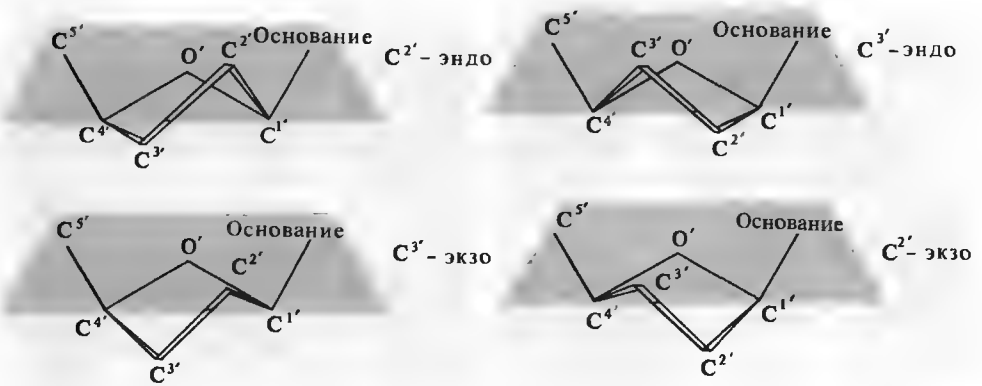


РИС.3.15. Четыре возможные конформации пиранозного кольца рибозы и дезоксирибозы. Четыре атома ($C^{1'}$, $C^{4'}$, O′ и либо $C^{2'}$, либо $C^{3'}$) расположены почти в одной плоскости. В ДНК-В остаток сахара находится в $C^{2'}$-*эндо*-конформации, а в ДНК-А и РНК-11 — в $C^{3'}$-*эндо*-конформации. [V.A. Bloomfield, D.M. Crothers, I. Tinoco, Jr., Physical Chemistry of Nucleic Acids (New York: Harper and Row, 1974).]

ческой невыгодности перехода всех сахаров в $C^{2'}$-*эндо*-конформацию и из-за стерических препятствий, которые вызвала бы в структуре эта конформация сахара.

Некоторые физические свойства ДНК в растворе плавно меняются при небольших изменениях внешних условий. Эти изменения не связаны ни с разрывом водородных связей, ни с другими сопряженными с денатурацией спирали явлениями. Они отражают скорее всего сдвиг равновесной конформации двойной спирали внутри данного структурного семейства при изменениях концентрации соли или температуры. Для некоторых регулярных последовательностей, таких, как синтетические полинуклеотиды, содержащие только dA в одной цепи и только dT — в другой, определенные структуры оказываются более предпочтительными. Возможно, наличие таких последовательностей в природных ДНК приводит к локальным изменениям структуры. Это могло бы объяснить, как белки, связывающиеся с ДНК, узнают определенные участки двойной спирали без разделения цепей, необходимого для взаимодействия с основаниями.

Случаи одновременного существования А- и В-форм в пределах одной двухцепочечной молекулы пока не обнаружены. Основываясь на изменении в угле наклона пары к оси спирали, Страсер и Арнотт указали на то, что в месте стыка А- и В-спиралей должен образовываться изгиб, если сохраняется непрерывность стэкинга оснований в этой точке. С другой стороны, была выдвинута гипотеза, что при синтезе ДНК могут использоваться РНК-затравки. В таком случае на начальных стадиях синтеза ДНК должна реализовываться следующая структура:

	3′	dNpdNpdNpdNpdNpdNpdNpdNp
	5′	**rMprMprM**pdMpdMpdM
Предпочтительная форма спирали		A i B

M — основание, комплементарное N; d и r соответствуют дезоксирибозе и рибозе; i — граница раздела А- и В-спиралей — предпочтительных форм. Здесь гибридная спираль РНК — ДНК (образующаяся из-за наличия РНК-затравки) переходит в чисто ДНК-спираль, образованную новосинтезированной ДНК. Важно знать, происходит ли излом на границе, и если нет, то как далеко по спирали распространяется возмущение, обуслов-

ленное наличием фрагмента, находящегося в А-форме. Исследования модельных последовательностей позволяют предположить, что переход в А-форму одной области спирали может влиять на участки, отстоящие от этой области на значительное число пар.

Тот факт, что ДНК существует как в А-, так и в В-формах, а РНК только в одной из форм А-семейства, может иметь важное биологическое значение. В принципе это открывает путь к созданию соединений — в природе или в лаборатории, — которые узнают или обе молекулы в двухцепочечной форме, или только двойную спираль В-формы ДНК. Взаимодействия между этими соединениями и нуклеиновыми кислотами могли бы не зависеть от последовательности, поскольку А- и В-формы различаются довольно сильно. Для того чтобы распознать таким же общим способом РНК, можно попытаться найти одноцепочечные области или петли в шпильках, поскольку эти элементы являются общими для всех известных природных РНК и, по-видимому, отсутствуют в ДНК на всех стадиях ее функционирования, за исключением, возможно, репликации.

ДРУГИЕ ПОЛИНУКЛЕОТИДНЫЕ СПИРАЛИ

Известно несколько двойных или множественных полинуклеотидных спиралей с параллельными цепями. В таких структурах уотсон-криковское спаривание оснований не сохраняется. Наиболее изучены двухцепочечные полимеры, в которых обе цепи представляют собой гомополинуклеотиды одного и того же основания. В табл. 3.1 перечислены некоторые структурные особенности двойных спиралей, одна из которых образована протонированной poly(A), а другая — наполовину протонированной poly(C). Легко видеть, что геометрия этих спиралей совершенно другая, чем у спиралей А- и В-семейств.

Строение пар оснований в случае протонированных полинуклеотидов рассмотрено в гл. 22. Эти пары и соответствующие спирали имеют оси симметрии C_2, параллельные спиральным осям. Poly (I) образуют тройную спираль с осями C_3, также параллельными спиральным осям. Такая симметрия возможна только для параллельных полинуклеотидных цепей. Известно множество других трехцепочечных структур, в которых дополнительная цепь образует пары оснований с одной или двумя обычными цепями антипараллельной уотсон-криковской спирали. Такие структуры наблюдаются главным образом в гомополимерных комплексах. Схемы спаривания оснований в двух трехцепочечных структурах приведены на рис. 22.14.

Строение одной из тройных спиралей, poly (dT) — poly (dA) — poly (dT), представлено на рис. 3.16. Poly (dA) и одна из poly (dT)-цепей образуют обычную двойную спираль со спаренными основаниями, с тем исключением, что ее геометрия ближе к ДНК-А, чем к более распространенной ДНК-В. Вторая dT-цепь прекрасно укладывается в большой желобок спирали А-типа, образуя дополнительные пары с NH_2-группой и атомом N^7 аденина.

Вообще говоря, образование трехцепочечных структур для большинства последовательностей кажется маловероятным. Во всяком случае, неизвестно, как должны быть расположены цепи с произвольными последовательностями, чтобы образовалась непрерывная симметричная тройная спираль. Короткие трехцепочечные участки могут формироваться в определенных областях РНК. В качестве примера можно привести трехцепочечную структуру $A \cdot U_2$ (рис. 22.12), очень стабильную при высокой ионной силе или высокой концентрации Mg^{2+}, т.е. в условиях, благоприятных для формирования прочных структур молекулами тРНК, рРНК, вирусными РНК. Тем не менее тройные спирали могут образоваться лишь при наличии подходящих последовательностей и при весьма изощренной укладке одиночной цепи. Как мы увидим ниже, тройные пары оснований были обнаружены в тРНК.

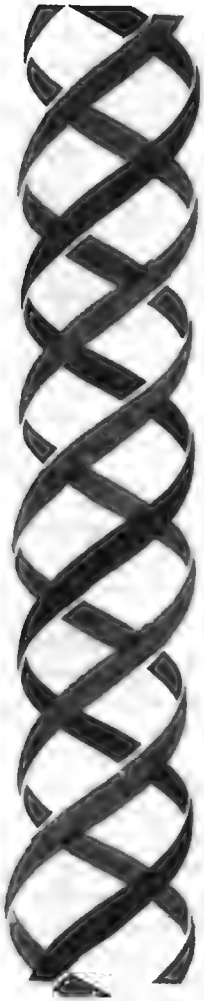


РИС.3.16. Тройная спираль poly (dT) — poly (dA) — poly (dT) по данным рентгеноструктурного анализа волокон. Poly (dA) и одна из цепей poly (dT) образуют такую же спираль, как и ДНК-А. Третья цепь, poly (dT) (выделенная на рисунке цветом), лежит в большом желобке этой спирали [S. Arnott et al., Nucleic Acids Res., 3, 2459 (1976).]

3.5. Третичная структура

ТРЕТИЧНЫЕ ВЗАИМОДЕЙСТВИЯ В тРНК С УЧАСТИЕМ ВОДОРОДНЫХ СВЯЗЕЙ

В настоящее время установлена детальная трехмерная структура только для одной РНК — дрожжевой фенилаланиновой тРНК; мы детально рассмотрим ее в этой и последующей главах. Хотя очень трудно не поддаться соблазну рассматривать целый ряд интересных особенностей $тРНК^{Phe}$ как отражение общих закономерностей строения таких молекул, не следует делать слишком общих выводов на основании анализа этой единственной известиой третичной структуры РНК (например, третичная структура миоглобина — первого исследованного в этом отношении белка — выглядит в ретроспективе нетипичной во многих отношениях).

Третичную структуру $тРНК^{Phe}$ независимо определили несколько групп исследователей методом рентгеноструктурного анализа. Были исследованы две кристаллические

формы, различающиеся симметрией упаковки молекул тРНК^{Phe}. Третичные структуры, соответствующие этим формам; почти идентичны, за исключением положения CpCpA-конца (рис. 3.17). Это является веским аргументом в пользу того, что кристаллические структуры практически не искажеиы межмолекулярными силами и достаточно близки по своей коиформации к изолированиым молекулам в растворе.

Прежде всего рассмотрим главные особенности структуры тРНК^{Phe}. По своей форме молекула похожа на букву Г. Поперечный размер отдельных участков не превышает диаметра двойной спирали ДНК. Структура намного более вытянута, чем в случае типичных глобулярных белков соответствующей молекулярной массы. В основе трехмерной конформации тРНК^{Phe} лежит вторичная структура «клеверный лист» (рис. 3.17,А), стабилизированная большим числом дополиительных парных взаимодействий между основаниями. Поскольку эти взаимодействия затрагивают удаленные друг от друга участки молекулы, образующиеся пары оснований называются третичными.

Стебли клеверного листа уложены в третичной структуре таким образом, что образуют две более длинные спиральные области: спираль акцепторного стебля непрерывно переходит в спираль псевдоуридилового стебля, образуя одии из «рукавов» Г-образной молекулы. Другой «рукав» образован антикодоновым и дигидроуридиловым стеблями, имеющими почти параллельиую ориентацию. Непрерывность между этими двумя спиральными участками обеспечивается иеобычной парой осиований, m^2G^{26} — A^{44} (рис. 3.18,*А*). Третичные взаимодействия удлиняют оба рукава буквы Г. Так, одна дополнительная пара образуется внутри псевдоуридиловой петли и две — между псевдоуридиловой и дигидроуридиловой петлями; в итоге длина псевдоуридилового стебля увеличивается на три пары оснований.

Наиболее иитересные особенности третичной структуры тРНК связаны со строением увеличенной D-спирали (рис. 3.17). Здесь и только здесь в молекуле образуется комплекс из многих звеньев — элемент третичной структуры, характерный для белков. Увеличенная D-спираль удлиняет и продолжает спираль дигидроуридиловой шпильки. Здесь, однако, образуется четырехцепочечная структура, включающая основания дигидроуридиловой петли (1), дополнительной петли (2), два основания между акцепторным и дигидроуридиловым стеблями, занимающие положения 8 и 9 (иногда называемые мини-петлей) (3), и две цепи дигидроуридилового стебля (4). Увеличенная D-спираль содержит несколько троек оснований. Например, гуанин в положеиии 22 образует водородные связи с двумя другими основаииями, C^{13} и m^7G^{46} (рис. 3.18,*Б*). Отметим, однако, что эти тройки не образуют тройных спиралей, поскольку включают разные комбинации цепей.

Существует шесть различных способов объединения в пары четырех цепей. Тройки оснований и третичные пары в увеличеиной D-спирали реализуют пять из этих шести возможностей. Это означает, что для рассматриваемой области характерна очень плотная упаковка и, вероятно, ее структура является весьма жесткой.

Другой иеобычной особенностью третичной структуры тРНК^{Phe} является наличие большого числа водородных связей между основаниями и фосфодиэфириым остовом. Укажем иа взаимодействие между фосфатом остатка A^9 и аминогруппами остатков C^{13} и m^7G^{46} (рис. 3.18,*Б*), а также на водородную связь между 2′-ОН в рибозе остатка U^8 и атомом N^1 остатка A^{21} (рис. 3.18,*В*). В третичной структуре имеется по крайней мере пять таких взаимодействий между основаииями и 2′-группами рибозы. Эти связи, по-видимому, играют существеииую роль в стабилизации третичной структуры. Более того, обнаружение таких взаимодействий позволяет выдвинуть гипотезу о том, почему в природе используются молекулы как ДНК, так и РНК. Остов ДНК гораздо более стабилен химически, чем остов РНК, но из-за отсутствия 2′-ОН-групп он, вероятно, в меньшей степени способен к формированию компактной третичной структуры.

Еще одна интересная особенность структуры тРНК^{Phe} состоит в иаличии нескольких резких изгибов фосфодиэфириой цепи. Как будет показано в гл. 6, остов нуклеиновой кис-

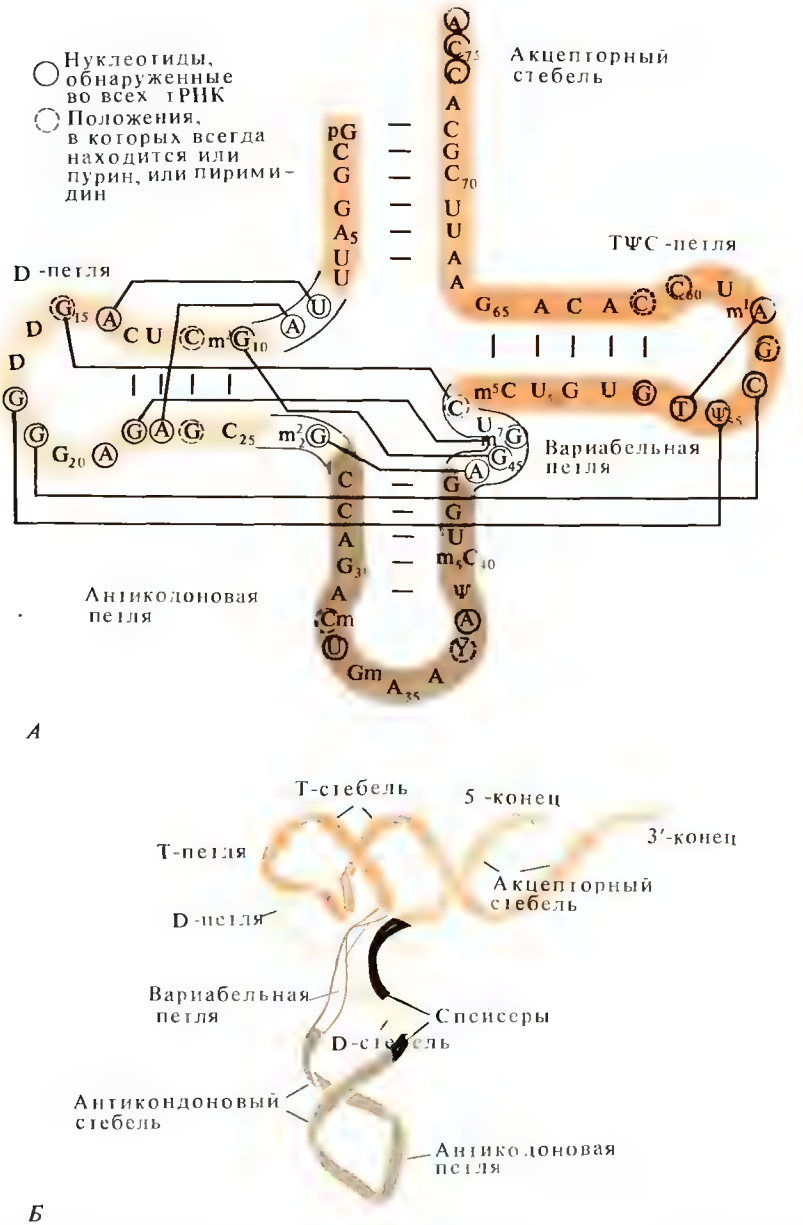
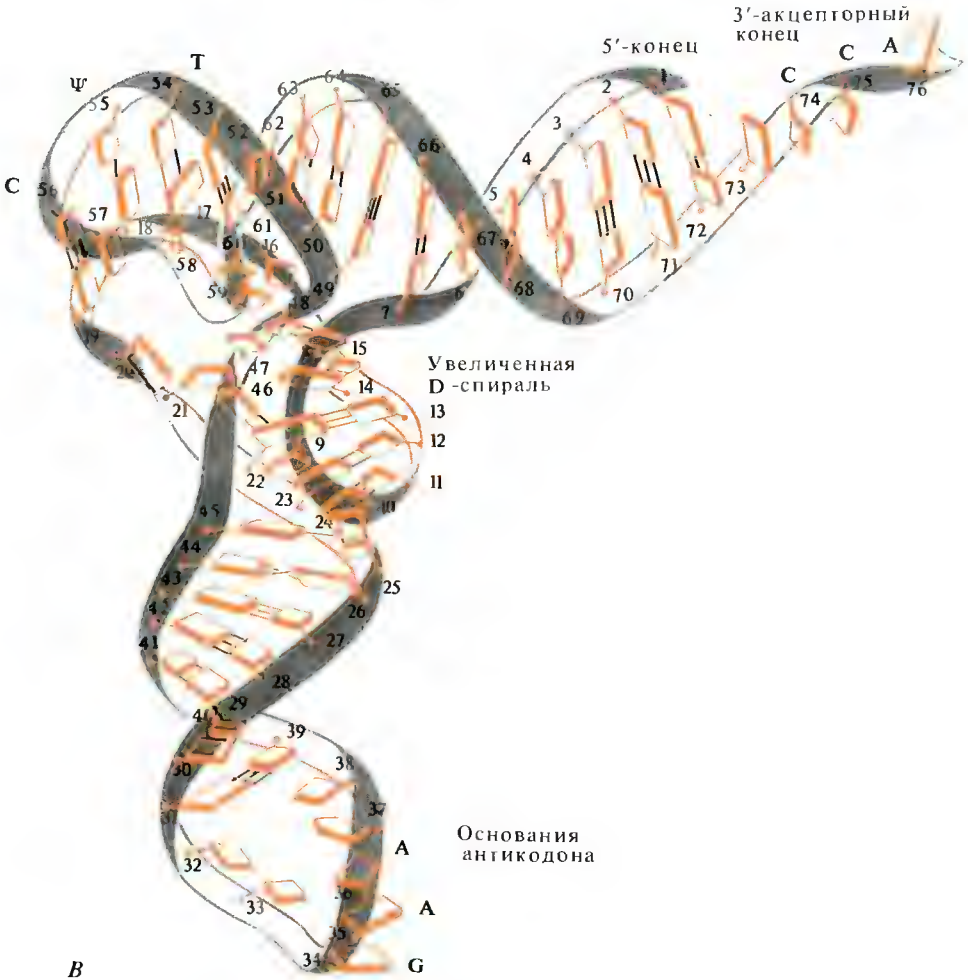


РИС.3.17. Третичная структура дрожжевой тРНК^{Phe}. *А*. Нуклеотидная последовательность. Отмечены неизменные для всех первичных структур остатки, а также положения, в которых всегда находится либо пурин, либо пиримидин. Остатки, взаимодействующие на уровне третичной структуры, соединены. [G.J. Quigley, A. Rich, Science, **194**, 791 (1976).] *Б*. Остов орторомбической формы по С. Киму, А. Ричу и др. Структура моноклинной формы, определенная Дж. Робертсом, А. Клугом и др., почти не отличается от приведенной. Различные петли и стебли для удобства идентификации окрашены на рис. *А* и *Б* в соответствующие цвета. *В*. Полная третичная структура. Пурины изображены в виде вытянутых брусков, пиримидины — в виде квадратных брусков; линии между брусками — водородные связи. (Рисунок Ирвинга Гейса.)



лоты является довольно жестким, поскольку лишь ограниченный диапазон значений углов вращения сахарофосфатной цепи благоприятен с энергетической точки зрения. Тем не менее третичные водородные связи, как показано на рис. 3.19, по-видимому, стабилизируют образование изгиба остова в антикодоновой и псевдоуридиловой петлях. Этот изгиб был назван уридиновым, поскольку основания U или Ψ играют ключевую роль в образовании этих третичных водородных связей. Возможно, уридиновый изгиб играет такую же роль в третичной структуре нуклеиновых кислот, как и β-изгиб в белках.

Вторичная структура тРНК^{Phe}, соответствующая модели клеверного листа, содержит 20 пар оснований, образующих 52 водородные связи. В третичной структуре к ним добавляется по крайней мере 40 дополнительных водородных связей. Большая часть из них не соответствует обычной уотсон-криковской схеме спаривания оснований. Важная роль третичных взаимодействий подтверждается тем фактом, что они всегда затрагивают или неизменные для всех известных последовательностей звеньев, или звеньев, которые меняются координированно в различных последовательностях, сохраняя возможность третичного спаривания. Таким образом, имеются веские основания считать, что главные

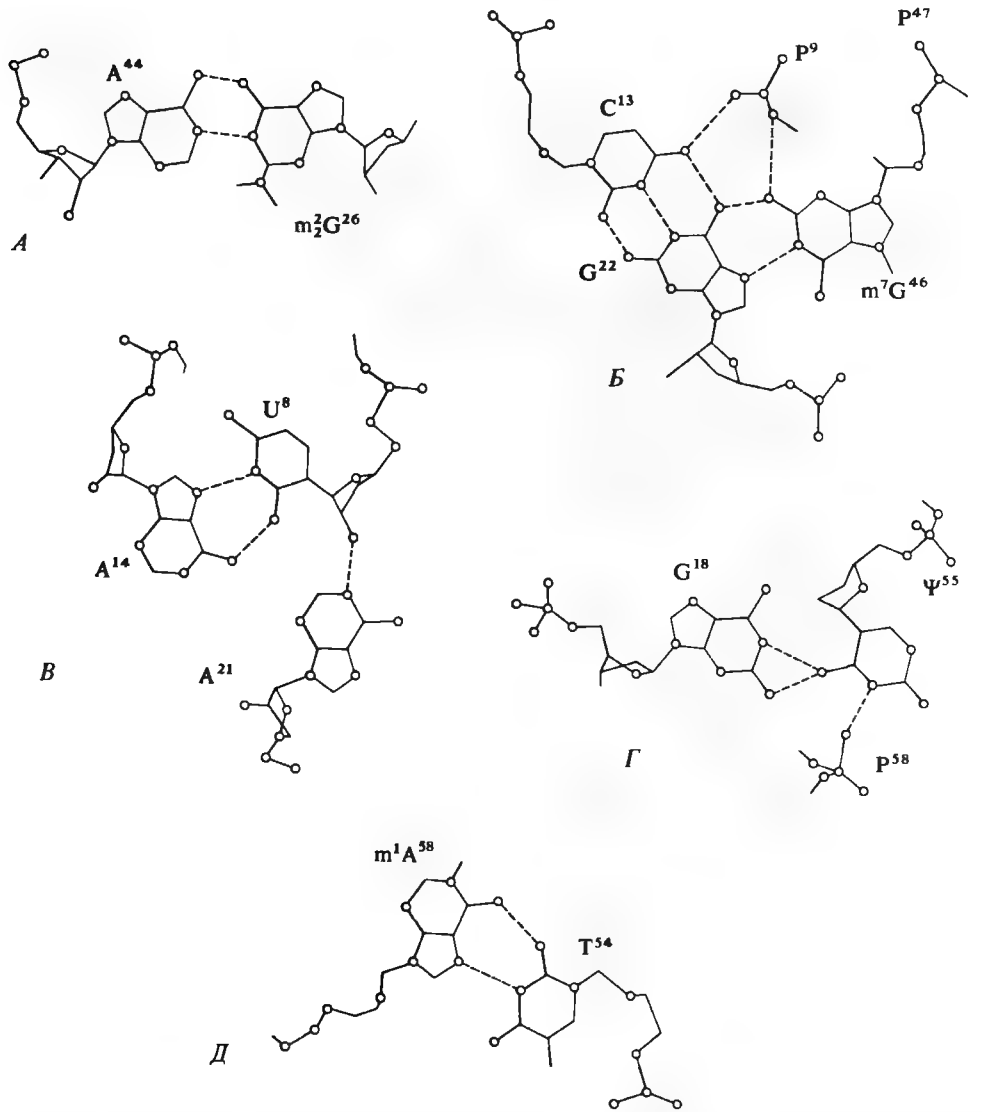


РИС.3.18. Некоторые третичные водородные связи в дрожжевой тРНК^{Phe}. Кружки — атомы кислорода, азота и фосфора; пунктирные линии — водородные связи (см. текст.). [G.J. Quigley, A. Rich, Science, 194, 791 (1976).]

черты третичной структуры тРНК являются общими для всех видов живых организмов. Схемы некоторых взаимодействий такого рода, реализующихся в дрожжевой тРНК^{Phe} и постулированных для других тРНК, приведены на рис. 3.20.

Предположение об образовании GC-пары между 15-м и 48-м звеньями было высказано М. Левиттом. Он обратил внимание на то, что если в положении 15 находится гуанин, то в положении 48 — цитозин, а если в положении 15 стоит аденин, то в положении 48 — урацил. Однако две цепи, содержащие эти основания, ориентированы параллель-

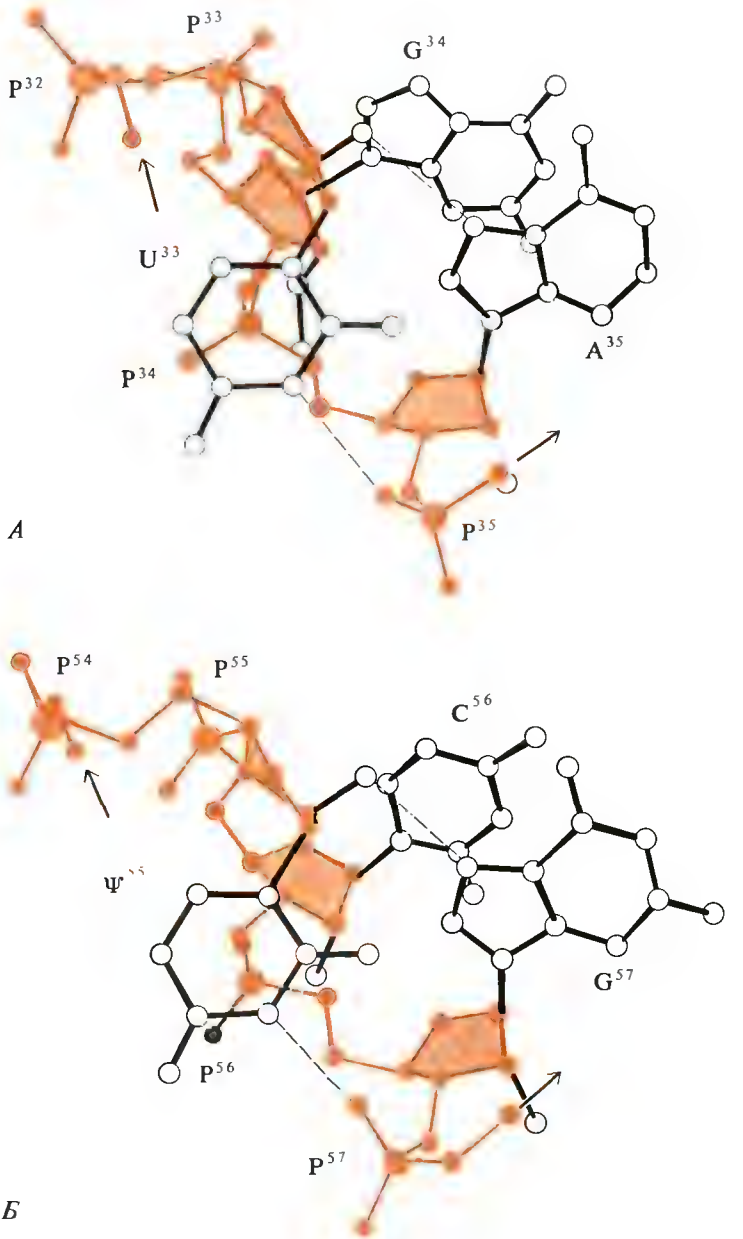
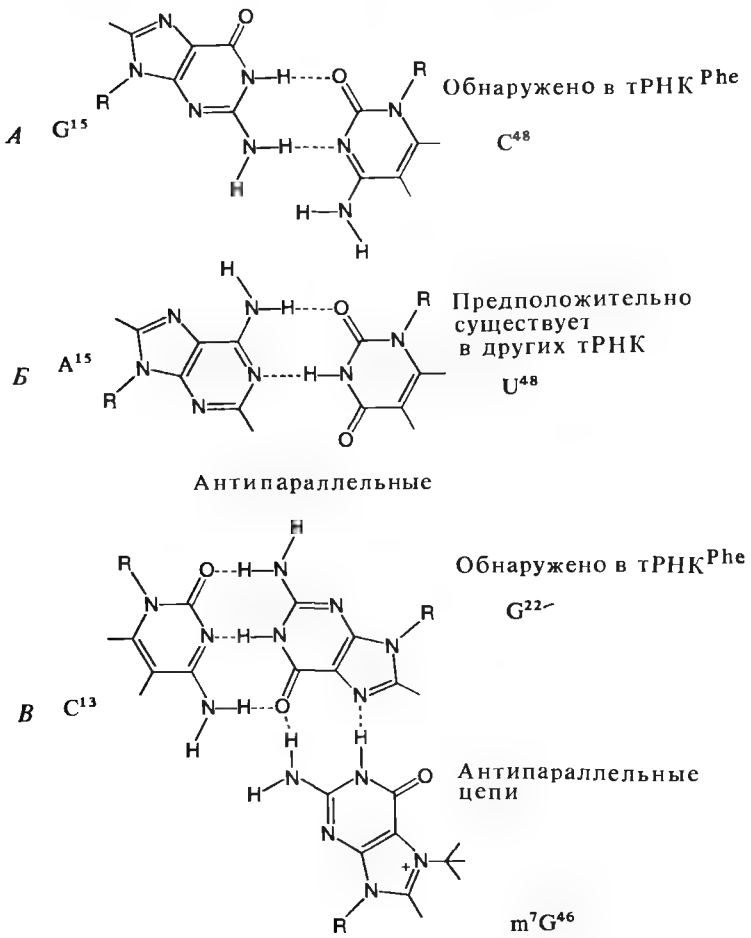


РИС.3.19. Два примера уридиновых изгибов (см. текст) в тРНК^{Phe}. Фосфодиэфирный остов и пентозное кольцо выделены цветом. Фосфаты перенумерованы, как 3′-эфиры нуклеотидов. Указаны водородные связи между 2′-OH-группой остатка U^{33} или Ψ^{55} и атомом N^7 остатка A^{35} или G^{57}, а также между (N^3 — H)-группой остатка U^{33} или Ψ^{55} и фосфатом A^{35} или G^{57}. Эти третичные взаимодействия, по-видимому, фиксируют поворот цепи. [G.J. Quigley, A. Rich., Science, **194**, 791 (1976).]

но. Поэтому не удивительно, что образовавшаяся пара не является уотсон-криковской: одно из оснований в ней повернуто приблизительно на 180° по отношению к его положению в уотсон-криковской паре. Получающаяся структура может быть образована как GC-, так и AU-парой (рис. 3.20,*А*). В этих парах, однако, не существует оси симметрии 2-го порядка и, следовательно, образование CG- или UA-пар должно быть сопряжено с изменением фосфодиэфирного остова. Но поскольку такие комбинации никогда не наблюдались, структура остова в этой области остается, по-видимому, неизменной.

В качестве второго примера рассмотрим тройку оснований C^{13} — G^{22} — m^7G^{46} (рис. 3.20,*В*). В этом случае первые два основания образуют уотсон-криковскую пару. Эти три основания в указанных положениях присутствуют в большинстве других видов тРНК из этого же класса. В одном случае, $тРНК^{Arg}$ из дрожжей, они заменены на U^{13} — A^{22} — A^{46}, однако такую последовательность можно рассматривать как тройку с эквивалентной структурой. В другом случае мы имеем тройку U^{13} — G^{22} — G^{46}, и здесь возможна та же структура, если допустить образование GU-пары между первыми двумя основаниями (рис. 3.20,*Г*,*Д*).



Основания U^8, A^{14} и A^{21} присутствуют во всех тРНК, и, следовательно, комплекс, который они образуют, тоже может составлять важный структурный элемент этих молекул (рис. 3.18,*В*). Аналогичным образом последовательность G^{18} — G^{19}, как и последовательность Ψ^{55} — C^{56}, обнаружена во всех тРНК; таким образом, третичное взаимодействие между этими двумя парами имеет место, по-видимому, во всех тРНК.

Структурным элементом, характерным для всех тРНК, должна быть и пара оснований T^{54} — m^1A^{58} (рис. 3.18,*Д*). Отметим, что пара оснований, включающая m^1A, не может быть уотсон-криковской, поскольку азот N^1, являющийся акцептором водорода в обычной АТ-паре, несет метильную группу. В структуре $тРНК^{Phe}$ акцептором в паре T^{54} — m^1A^{58} служит азот N^7 аденина. Такой тип спаривания оснований называется хугстеновским, по имени кристаллографа, который первым обнаружил его в комплексах компонентов нуклеиновых кислот.

Есть основания ожидать, что все тРНК — не только те из них, которые относятся к тому же классу, что и дрожжевая $тРНК^{Phe}$, — имеют сходную третичную структуру. Молекулы тРНК других классов могут иметь большую дополнительную петлю, у них может

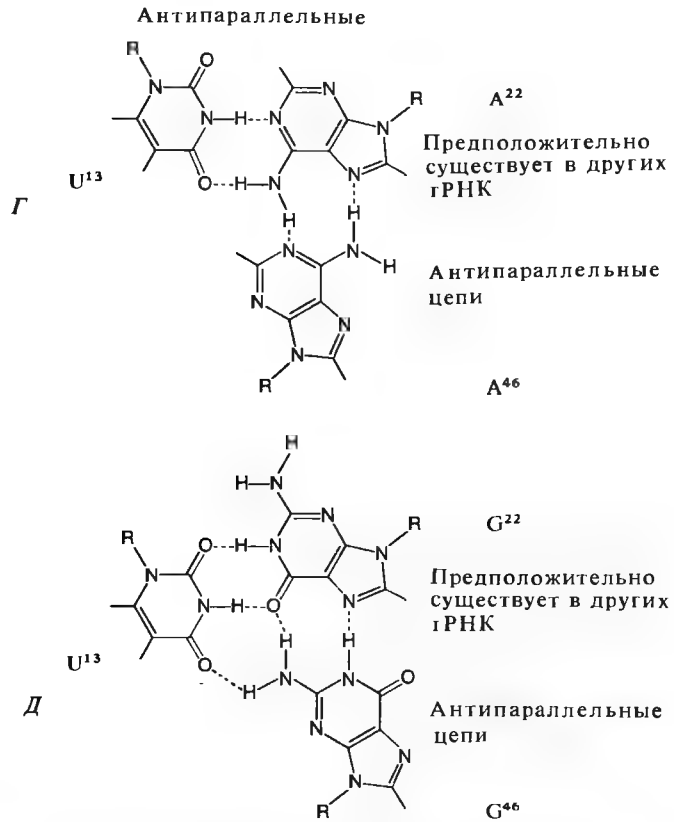


РИС.3.20. Два третичных взаимодействия между основаниями в дрожжевой тРНК и похожие взаимодействия, предположительно существующие в других тРНК (см. текст). [A. Klug et al., J.Mol.Biol., **89**, 511 (1974).]

быть другой размер дигидроуридиловой петли и т.д. Однако эти изменения всегда затрагивают остатки, находящиеся на поверхности структуры, и поэтому дополнительные остатки могут быть встроены в молекулу без радикальной перестройки ее третичной структуры.

СТЭКИНГ ОСНОВАНИЙ В тРНК

До сих пор, рассматривая тРНК^{Phe}, мы говорили только об образовании в ней пар оснований, теперь же остановимся на другом очень важном аспекте этой структуры. Известно, что между плоскими ароматическими соединеииями, такими, как пурины или пиримидины, может существовать сильное межплоскостное взаимодействие (стэкинг), способствующее образованию стопкообразных структур. Такие структуры наблюдаются как в кристаллах, так и в растворе. Термодинамическая природа стэкинга подробно обсуждается в гл. 6. Здесь достаточно отметить, что межплоскостное взаимодействие пар хотя и лишено специфичности по отношению к типам оснований, играет не меньшую роль в стабилизации конформации нуклеиновых кислот, чем спаривание оснований. Почти за десять лет до установления третичной структуры тРНК^{Phe} было высказано утверждение, что в структуре тРНК должно быть максимально возможное число межплоскостных взаимодействий. Это предсказание полностью подтвердилось для тРНК^{Phe} (рис. 3.17). Только четыре из 76 оснований в молекуле (U^{47}, D^{16}, D^{17} и G^{20}) не участвуют в стэкинге. Дигидроуридин вообще не может образовывать стопок, поскольку восстановленное пиримидиновое кольцо уже не является плоским. Поэтому можно сказать, что стэкинг в структуре тРНК затрагивает почти максимально возможное число оснований. Из рис. 3.17 видно, что все основания, вовлеченные в стэкинг, образуют две непрерывные взаимно перпендикулярные структурные единицы. Все точки цепи тРНК, в которых имеется структурная гибкость, локализованы, по-видимому, в областях, где нарушается непрерывность стэкинга.

Интересной особенностью стэкинга в тРНК является интеркаляция некоторых оснований между другими. Как можно заметить из рис. 3.17, основания U^8 и A^9 сильно раздвинуты, так что расстояние между ними в вертикальном направлении составляет около 6,5 Å вместо обычных 3,0 Å, и между этими двумя основаниями расположено основание C^{13}, которое, следовательно, взаимодействует с каждым из них.

ТРЕТИЧНАЯ СТРУКТУРА тРНК^{Phe}; БИОХИМИЧЕСКИЕ И ХИМИЧЕСКИЕ СВОЙСТВА ЭТОЙ МОЛЕКУЛЫ

Стэкинг и спаривание оснований в тРНК^{Phe} приводят к тому, что подавляющее большинство оснований оказываются недоступными для растворителя. Наиболее существенными исключениями являются CpCpA-конец и три основания, составляющие антикодон. Тот факт, что эти основания не маскированы, т.е. доступны для взаимодействий, вполне понятен: последние должны участвовать в спаривании с основаниями мРНК на рибосоме, первые же почти наверняка участвуют в связывании тРНК с пептидил-трансферазным центром 50S-субчастицы рибосомы, поскольку даже фрагменты тРНК, содержащие просто CpCpA-аминоацил, также могут успешно связываться и функционировать при определенных условиях.

Довольно давно известно, что большинство оснований тРНК оказывается защищено от действия различных химических реагентов. Эта особенность тРНК явилась одним из первых свидетельств наличия у этой молекулы определенной третичной структуры. Отталкиваясь от известной третичной структуры, можно попытаться проанализировать, как она согласуется с данными о реакционной способности отдельных оснований. На рис. 3.21 указаны звенья дрожжевой тРНК^{Phe}, которые взаимодействуют с реагентами,

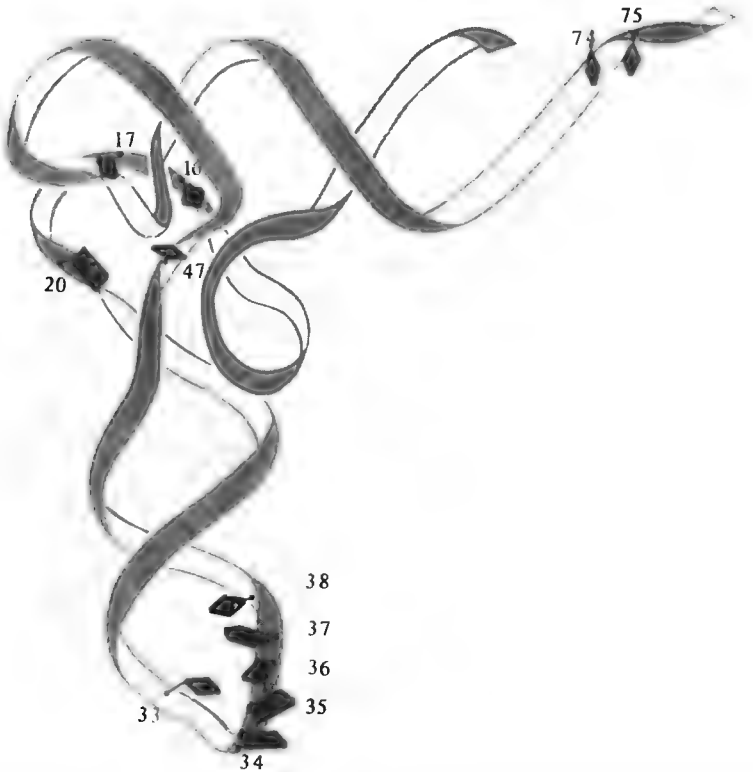


РИС.3.21. Химическая модификация дрожжевой тРНК^{Phe} реагентами, специфически взаимодействующими лишь с неспаренными основаниями. [Основная часть данных из работы D.Rhodes, J.Mol.Biol., **94**, 449 (1975) предоставлена Кларком. Рисунок Ирвинга Гейса.]

специфичными к цитидину (метоксиамин), гуанину, урацилу и дигидроуридину (карбодиимид), аденину (надфталевая кислота) и гуанину (кетоксаль). Ясно, что лишь несколько оснований легко доступны для действия этого широкого спектра реагентов, и высокую реакционную способность этих оснований можно понять исходя из известной трехмерной структуры тРНК. Большинство экспериментов по химической модификации других тРНК удается объяснить, предполагая, что все молекулы тРНК имеют одинаковую третичную структуру.

У многих типов тРНК *E. coli* в 8-м положении вместо обычного урацила, как у дрожжевой тРНК^{Phe}, присутствует 4-тиоуридин. Известно, что соответствующий атом серы может быть модифицирован даже сравнительно объемистыми реагентами без какого-либо изменения способности тРНК участвовать в различных этапах белкового синтеза. Для этого имеются две причины. Во-первых, U^8 в тРНК^{Phe} локализован у дна широкого желоба на поверхности молекулы. Во-вторых, хотя U^8 вовлечен во взаимодействие соответствующей тройки оснований, в образовании водородных связей участвует O^2, а не O^4 (см. рис. 3.18,*В*).

Особенно интересной химической модификацией, которая может происходить во многих тРНК из *E. coli*, является образование сшивки между 4-тио-U^8 и C^{13} под действием ультрафиолетового излучения; при этом тРНК полностью сохраняет свою биологическую активность. Отсюда следует, что основания 8 и 13 должны быть сближены в натив-

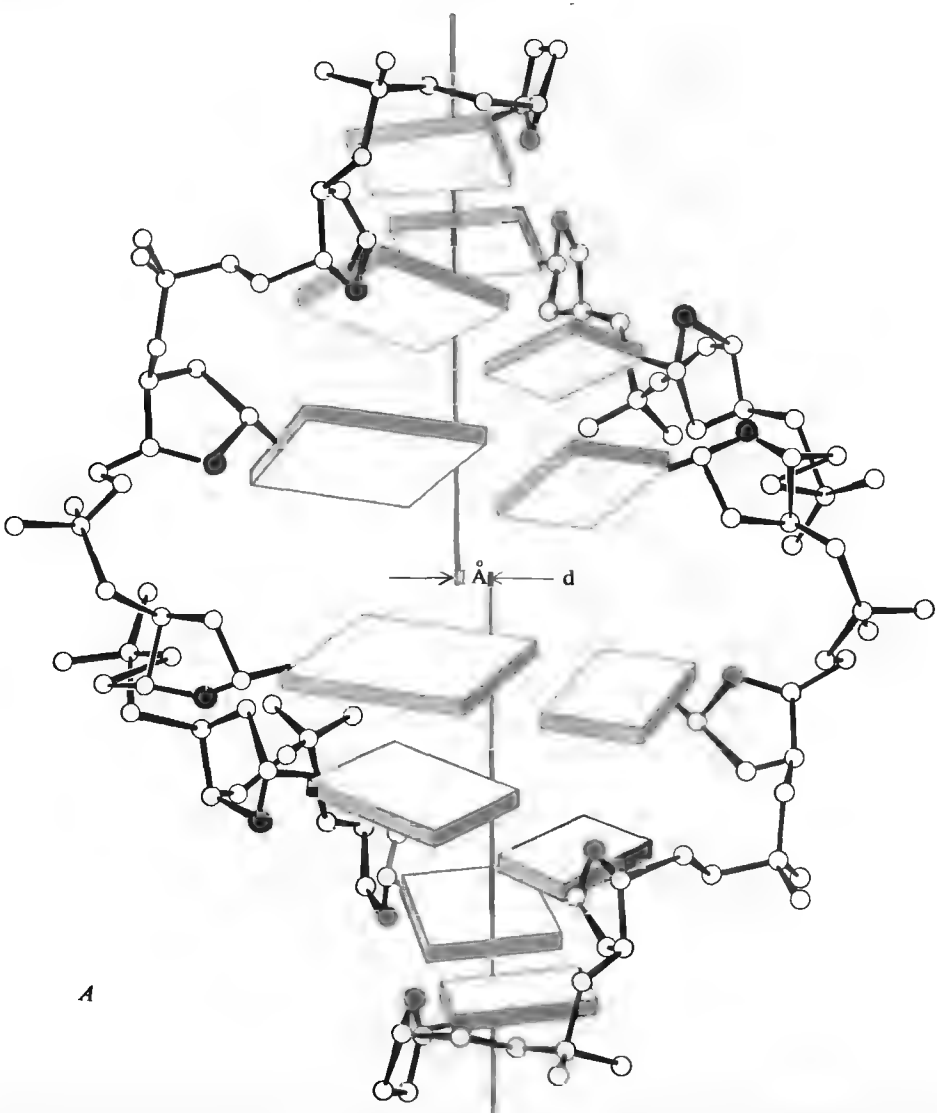
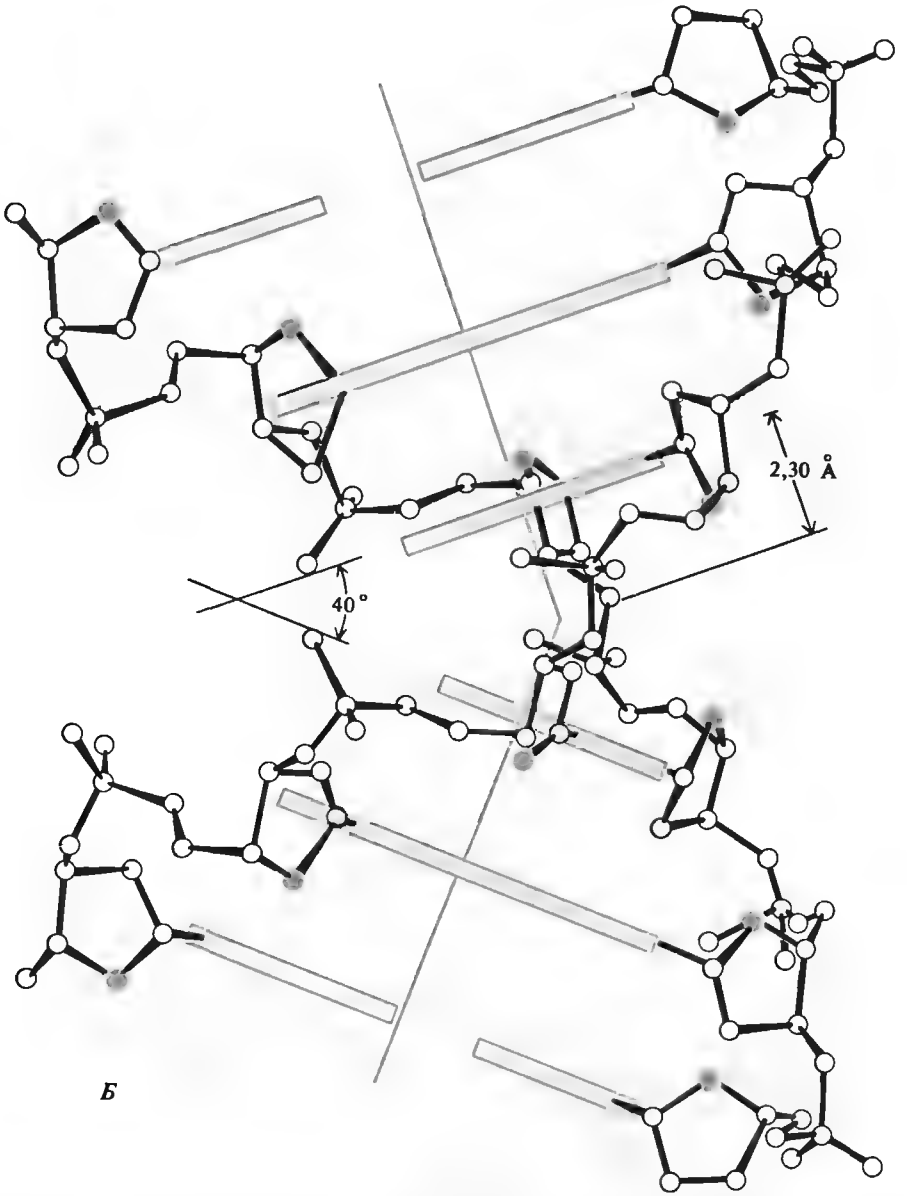


РИС.3.22. Структура возможных изломов в двухцепочечной ДНК. *А*. Вид в ваправлении оси симметрии 2-го порядка. Б. Вид сбоку. [H. Sobell et al., J.Mol.Biol., 114, 333 (1977).]

ной третичной структуре свободной молекулы — что и имеет место в действительности, как показано в предыдущих разделах. Эти основания должны оставаться сближенными на всех этапах функционирования молекулы. Фотохимия реакции между тио- U и цитидином была изучена на модельных системах, и, исходя из данных о механизме этой реакции и о структуре продукта, удалось точно предсказать, каково должно быть относительное пространственное расположение этих двух оснований в тРНК для эффективного образования фотосшивки. Это предсказание полностью подтверждается известной теперь пространственной структурой дрожжевой $тРНК^{Phe}$ (см. гл. 24).

Многие эксперименты по химической модификации тРНК были проведены до того, как с помощью рентгеноструктурного анализа удалось получить детальную информацию



Б

о ее пространственной структуре. Кроме того, всесторонние физико-химические исследования позволили сделать определенные выводы о спаривании оснований, о степени близости различных участков молекулы, об их общем размере и форме. Удалось установить последовательность оснований для более чем 30 разных тРНК, и было высказано предположение, что все эти тРНК имеют сходную структуру. Были выявлены некоторые мутации, приводящие к определенным функциональным изменениям, которые удалось связать с предполагаемыми изменениями в структуре. Правомерен вопрос: сколь успешны были попытки предсказать третичную структуру тРНК на основании совокупности данных, полученных для раствора? В определенной степени эти предсказания подтвердились. Так, Левитт высказал правильные предположения относительно половины третичных взаимо-

действий, и некоторые модельные структуры по своей форме были весьма близки к найденной. Тем не менее ни в одном случае не удалось правильно предсказать укладку полинуклеотидной цепи. Одна из трудностей состояла в том, что, исходя из совокупности имеющихся надежных данных, можно было построить довольно много разных моделей, а остальные данные были не столь бесспорны, чтобы на основании их можно было сделать выбор между этими моделями. Теперь, когда мы знаем третичную структуру одной РНК, было бы интересно выяснить, могут ли быть использованы принципы, выведенные из этой структуры, для более точного предсказания структуры других небольших РНК.

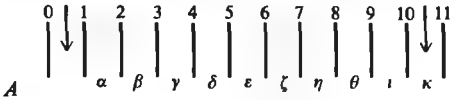
Хотя детальная информация по этому поводу на сегодняшний день отсутствует, кажется весьма вероятным, что другие небольшие РНК также имеют характерную третичную структуру. Например, при мягкой обработке 5S-рРНК переходит в одну из нескольких денатурированных форм. Эти формы могут быть отделены одна от другой и от нативной структуры. Более крупные РНК имеют четко выраженные признаки третичной структуры, но неизвестно, характерен ли для них только один тип нативной третичной структуры. Исследования с помощью сшивок указывают, что в 16S- и 23S-рРНК удаленные друг от друга вдоль цепи участки оказываются сближенными и образуют двойные спирали. Расположение сшивок слабо зависит от окружения, в котором находится РНК, и это позволяет считать, что выявляемые особенности третичной структуры свойственны всем молекулам.

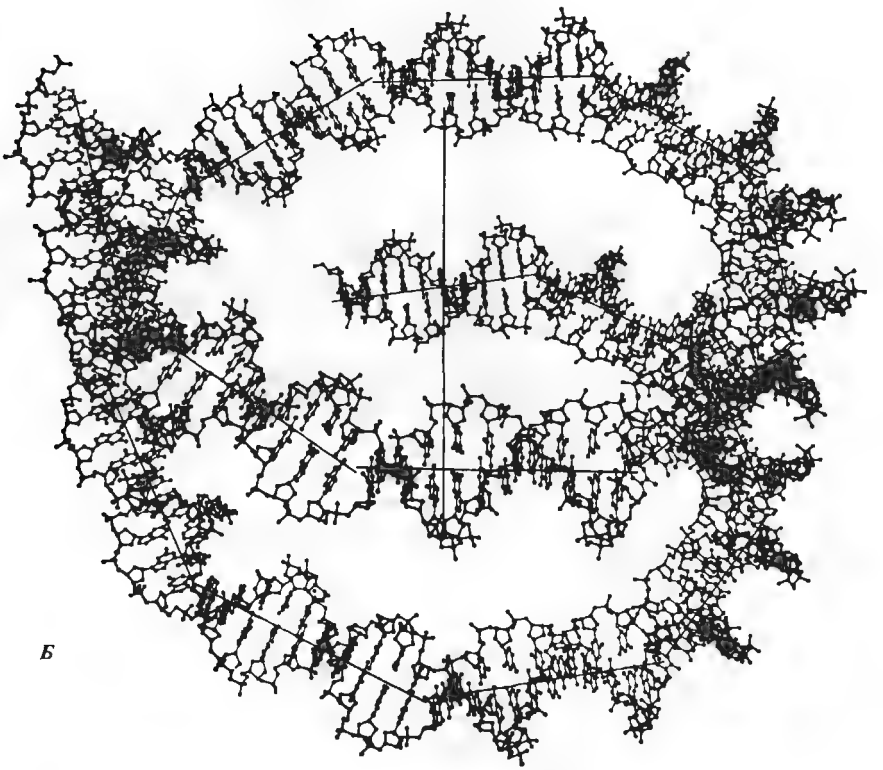
ТРЕТИЧНАЯ СТРУКТУРА ДНК

Детальная информация о третичной структуре ДНК в настоящее время отсутствует, однако существуют интересные гипотезы о некоторых возможных способах укладки этих молекул. Известно, что в клетках и вирусах ДНК обычно упакована в довольно компактные структуры. Наиболее интересен вопрос о том, как жесткая двойная спираль ДНК может изгибаться и складываться таким упорядоченным образом. Для объяснения гидродинамических свойств высокомолекулярной ДНК необходимо допустить, что двойная спираль обладает определенной гибкостью (см. гл. 19). ДНК моделируют персистентной, «червеобразной» цепью, которая может слегка деформироваться в каждом звене.

Такой макроскопический подход достаточен для объяснения гидродинамического поведения, но хотелось бы знать механизм гибкости на молекулярном уровне. Согласно идее, впервые высказанной Ф. Криком и А. Клугом, в ДНК могут образовываться изломы с нарушением стэкинга двух соседних пар оснований и небольшим раскручиванием остова. Г. Собелл подробно исследовал стереохимию возможных изломов на молекулярных моделях. Структура излома, образование которого кажется наиболее вероятным, приведена на рис. 3.22. В этом случае ось спирали изгибается на 40° и смещается на 1 Å. Пары оснований удалены на 2,30 Å от центра излома и смещены на 1,70 Å от оси псевдосимметрии C_2 в ДНК-В. Для того чтобы стал возможен этот излом, сахар в области излома должен перейти из конформации $C^{2'}$-*эндо* в ДНК-В в $C^{3'}$-*эндо* ($3' — 5'$)$C^{2'}$-*эндо*. Такое изменение может произойти также при раздвижке пар оснований, сопровождающей встраивание плоской молекулы между ними (гл. 23).

Существенной особенностью излома, показанного на рис. 3.22, так же как и других предложенных изломов, является то, что он может происходить только в одном направлении относительно остова. Таким образом, направление возможного излома будет поворачиваться вместе со спиралью при движении от некоторой фиксированной точки. Если изломы расположены вдоль спирали через одинаковые интервалы, они будут образовывать в ДНК сверхспиральную структуру. Обозначения для таких регулярно расположенных изломов приведены на рис. 3.23,*А*. Например, *к*-изломы расположены через каждые 10 пар оснований и порождают левую сверхспираль с диаметром около 100 Å, показанную на рис. 3.23,*Б*. Каждое звено сверхспирали состоит из 10 пар оснований. Сверхспираль делает приблизительно 1,5 витка на 140 пар оснований.





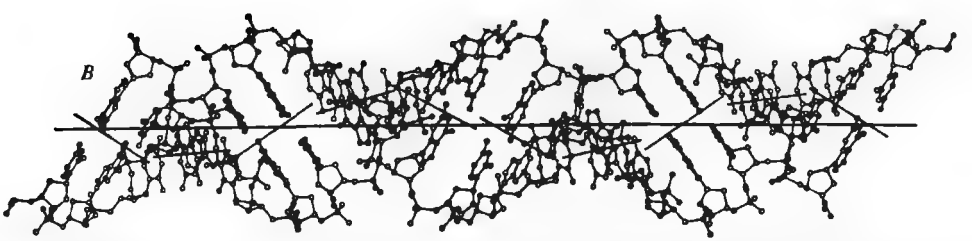


РИС.3.23. Сверхспирали, образованные регулярно расположенными вдоль цепи изломами. *А*. Обозначения регулярно расположенных изломов. *Б*. Двойная спираль ДНК, содержащая изломы через каждые 10 пар оснований (*κ*-изломы). *В*. Двойная спираль, содержащая изломы через каждые две пары (*β*-изломы). [H.Sobell et al., J.Mol.Biol., **114**, 333 (1977).]

Второй интересный пример сверхспирали, образованной изломами, приведен на рис. 3.23,*В*. Эта спираль содержит *β*-изломы. Она является правой и содержит 9,6 пары оснований на виток. В настоящее время, однако, нет прямых данных о существовании таких спиралей из изломов в упорядоченных структурах ДНК. Более того, многие исследователи отдают предпочтение модели, согласно которой изгибы в ДНК реализуются за счет

малых изменений углов вращения полинуклеотидного остова. Например, В. Олсон показала, что обширное семейство сверхспиралей ДНК может быть реализовано путем регулярной деформации спирали ДНК-В.

Более сложные примеры третичной структуры ДНК можно найти в случае кольцевых замкнутых двойных спиралей. Для третичной структуры таких молекул характерны весьма специфические особенности, обусловленные топологическими ограничениями в замкнутых цепях (гл. 24).

3.6. Четвертичная структура

Примеры четвертичной структуры в системах, состоящих исключительно из нуклеиновых кислот, очень немногочисленны. Один из них — четвертичная структура РНК опухолеродных вирусов, по-видимому, состоящей из двух одинаковых субъединиц. Электронно-микроскопические данные в пользу существования такой структуры в случае 52S-РНК из вируса RD-114 представлены на рис. 3.24. Стабилизация четвертичной структуры, вероятно, осуществляется благодаря спариванию оснований между двумя субъединицами РНК. Если это межцепочечное уотсон-криковское спаривание, то не совсем понятно, чем оно выгоднее внутрицепочечного.

Гораздо легче дать разумное объяснение образованию комплексов между неидентичными молекулами РНК. Например, молекулы тРНК с комплементарными антикодонами образуют очень прочный межмолекулярный комплекс in vitro. Роль этого явления in vivo неясна, но оно может служить указанием на то, что в РНК существуют какие-то не известные нам взаимодействия на уровне четвертичной структуры. Некоторая комплементарность существует между участками 16S- и 23S-рибосомной РНК из *E. coli*. Четвертичное спаривание оснований между этими участками может происходить, например, при образовании 70S-рибосомы из 30S- и 50S-субчастиц.

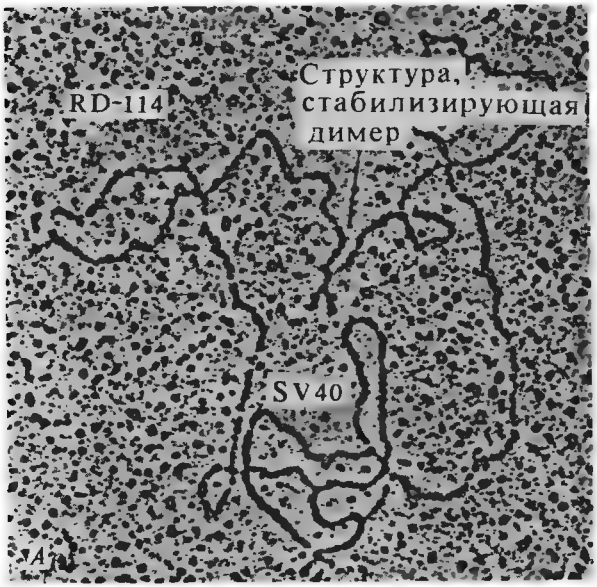
Краткие выводы

Нуклеиновые кислоты состоят из четырех нуклеозидов (A, G, C и T в ДНК или U в РНК), связанных фосфодиэфирными связями.

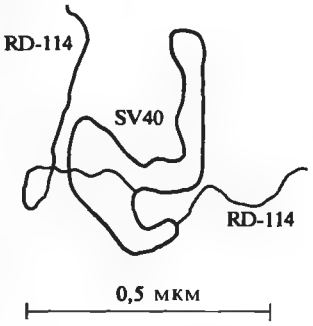
Нуклеотидный состав, в отличие от последовательности оснований, не играет существенной роли в определении структуры нуклеиновых кислот. Важнейшей чертой всех природных нуклеиновых кислот является сильное взаимодействие A с T или U и G с C, приводящее к образованию пар оснований. Оно играет решающую роль в образовании двухцепочечных спиральных структур, которые имеют оси симметрии второго порядка, лежащие в плоскости каждой пары оснований. Это означает, что две цепи двойной спирали должны иметь противоположную ориентацию. ДНК, как правило, состоит из двух полностью комплементарных цепей, образующих непрерывную двойную спираль. Молекула РНК обычно представляет собой одну ковалентную цепь, которая образует спирали, скручиваясь сама на себя и формируя таким образом определенную последовательность шпилек и петель.

Ограничениям, налагаемым на структуру спариванием оснований, удовлетворяет много различных спиральных конформаций. ДНК в водных растворах принимает В-форму двойной спирали, в которой пары оснований перпендикулярны оси спирали, но при определенных условиях она может быть переведена в А-спираль [1]. По своему строению А-

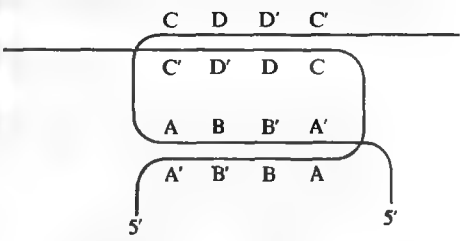
[1] Наряду с семействами правоспиральных В- и А-форм существуют участки ДНК, в которых чередуются пуриновые и пиримидиновые основания и которые могут переходить в левоспиральную Z-форму [см. A. H.-J. Wang et al., Nature, **282**, 680 (1979); A. Nordheim et al., Cold Spring Harbor Symp. Quant. Biol., **47**, 93 (1983) и другие статьи там же]. — *Прим. ред.*



А



Б



В

РИС.3.24. Комплексы РНК из опухолеродного вируса RD-114 и ДНК из SV40. RD-114 — эндогенный кошачий вирус типа С. Он содержит две субъединицы 26S-РНК, объединенные в одну структуру с константой седиментации 52S; каждая субъединица оканчивается сегментом poly A. Для того чтобы пометить концы цепей РНК, она была гибридизована с кольцевой ДНК SV40, модифицированной случайными вставками poly (dT). Структура образованного комплекса стабилизирована короткими (rA·dT)-участками. *А*. Электронная микрофотография. *Б*. Рисунок той части фотографии, на которой цепи RD-114 и SV40 недостаточно четко идентифицируются. Можно заметить, что две цепи РНК образуют сходные петли вторичной структуры на расстоянии около 2/3 от 3′-конца цепи. *В*. Возможная структура 5′-концов комплекса двух субъединиц РНК (действительная структура неизвестна). Буквы со штрихами обозначают участки, комплементарные тем, которые указаны соответствующими буквами без штрихов. [*А* и *Б* — из работы W. Bender, N. Davidson, Cell, 7, 595 (1976).]

форма довольно близка к спиралям РНК-11 или РНК-10, обнаруженным в двухцепочечной РНК. В этих формах пары оснований наклонены по отношению к оси спирали.

Единственной нуклеиновой кислотой с известной третичной структурой является дрожжевая $тРНК^{Phe}$. Ее структура стабилизирована большим числом водородных связей, отличных от тех, которые образуются при обычном спаривании оснований. Кроме того, в ней реализуется, по-видимому, максимально возможное число межплоскостных взаимодействий между соседними основаниями. Структура тРНК дает пример многоцепочечных взаимодействий и специфической конфигурации цепи, столь типичных для третичной структуры белков. По-видимому, важную роль в формировании третичной структуры РНК играет 2′-ОН-группа. Отсутствие этой группы в ДНК может объяснить, почему основные черты третичной структуры двух типов нуклеиновых кислот столь сильно различаются. Третичная структура ДНК изучена недостаточно, но может включать, по-

видимому, периодические или случайные изломы двойной спирали. Пока известно всего несколько примеров наличия четвертичной структуры у нуклеиновых кислот.

ЗАДАЧИ

3.1. Предположим, что мы хотим проверить, содержится ли в данной цепи ДНК из 10000 оснований определенная последовательность из 20 остатков. Для этого можно посмотреть, будет ли гибридизоваться комплементарная последовательность из 20 звеньев с этой ДНК. Однако синтезировать заданную последовательность из 20 оснований довольно трудно. Вычислите минимальную длину фрагмента этой последовательности, которую целесообразно было бы использовать в экспериментах по гибридизации. Эта длина должна быть достаточной, чтобы вероятность случайного появления комплементарного фрагмента в исследуемой ДНК не превышала 0,01. Можете считать, что ДНК содержит по 25% каждого основания.

3.2. ДНК имеет мол. массу $2{,}5 \cdot 10^8$. Вычислите ее длину в сантиметрах, предполагая, что ДНК находится в В-форме; в А-форме. Сколько витков содержит эта ДНК в каждой из форм?

3.3. Что произойдет, если смешать растворы двух комплементарных кольцевых одноцепочечных ДНК?

3.4. Предположим, что некий полинуклеотид предпочитает образовывать двойные спирали с параллельными цепями. Образование таких спиралей из длинных одноцепочечных молекул в принципе возможно. Как будет выглядеть такая структура?

3.5. Рассмотрим одноцепочечную ДНК с первичной структурой, в точности соответствующей первичной структуре $тРНК^{Phe}$. В чем будут состоять отличия третичной структуры, образуемой этой ДНК, от третичной структуры $тРНК^{Phe}$?

ЛИТЕРАТУРА

Общая

Bloomfield V. A., *Crothers D. M.*, *Tinoco I.*, Jr., 1974. Physical Chemistry of Nucleic Acids, New York, Harper and Row. [Монография.]

Cohn W. E., ed., 1976. Progress in Nucleic Acid Research and Molecular Biology, New York, Academic Press. [Сборникн обзорных статей; выходят примерно два тома в год.]

Duschensne J., ed., 1975. Physical Chemical Properties of Nucleic Acids, 3 vols., London, Academic Press.

Guschlbauer W., 1976. Nucleic Acid Structure, New York, Springer-Verlag.

Sundaralingam M., *Rao S. T.*, eds., 1975. Structure and Conformation of Nucleic Acids and Protein — Nucleic Acid Interactions, Baltimore, University Park Press.

Ts'o P. O. P., ed., 1975. Basic Principles of Nucleic Acid Chemistry, 2 vols., New York, Academic Press.

Специальная

Barrell B. G., *Clark B. F.*, 1974. Handbook of Nucleic Acid Sequences, Oxford, Joynson-Bruvvers.

Jack A., *Ladner J. E.*, *Klug A.*, 1976. Crystallographic refinement of yeast phenylalanine transfer RNA at 2,5 Å resolution, J. Mol. Biol., **108**, 619.

Kim S. H., 1978. Crystal structure of yeast $tRNA^{Phe}$. Its correlation to the solution structure and functional implications. In: Transfer RNA, ed. S. Altman, Cambridge, MIT Press, p. 248.

Ladner J. E., 1978. Structure of transfer ribonucleic acid. In: Biochemistry of Nucleic Acids, vol. 2, ed. B. T. C. Clark, Baltimore, University Park Press, p. 1.

Rich A., *Rajbhanday T.*, 1976. Transfer RNA: Molecular structure, sequence and properties, Ann. Rev. Biochem., **49**, 805.

Sobell H. M., *Tsai C. C.*, *Jain S. C.*, *Gilbert S. G.*, 1977. Visualization of drug—nucleic acid interactions at atomic resolution, III: Unifying structural concepts in understanding drug—DNA interactions and their broader implications in understanding protein—DNA interactions, J. Mol. Biol., **114**, 333.

Wu R., 1978. DNA sequence analysis, Ann. Rev. Biochem., **47**, 607.

Younghusband H. B., *Inman R. B.*, 1974. The electron microscopy of DNA, Ann. Rev. Biochem., **43**, 605.

Глава 4 Другие биологические полимеры

По существу все примеры, которые будут рассмотрены в последующих главах, основываются на исследовании чистых нуклеиновых кислот или белков. Эти биологические молекулы изучены к настоящему времени наиболее полно, поэтому на их примере удобнее всего показать, как существующие принципы и методы применяются для исследования структуры и функции биополимеров.

В данной главе будут описаны не столь хорошо изученные биологические системы. Они могут быть подразделены на три основных класса. К первому относятся биополимеры, состоящие из мономерных единиц, отличных от аминокислот или нуклеотидов; примером такого рода биополимеров служат полисахариды. Второй класс составляют биополимеры, содержащие более одного типа мономеров, или же комплексы из биологических макромолекул разных видов; это, например, нуклеопротеиды, гликопротеиды и гликолипиды. Третий класс включает ассоциаты макромолекул с участием липидов, такие, как липидный бислой клеточных мембран. Системы этого типа состоят из гетерогенных фаз, и поэтому их весьма трудно исследовать.

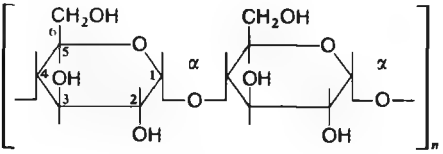
Сейчас пока неясно, могут ли существующие методы и подходы, в основном разработанные при исследовании белков и нуклеиновых кислот, полностью раскрыть поведение рассматриваемых здесь систем. Многие из этих систем являются предметом интенсивных исследований, и мы приведем основные сведения о них, полученные к настоящему времени, чтобы читатель мог сопоставить эти системы с белками и нуклеиновыми кислотами. (Вопросы, касающиеся мембран и взаимодействий между белками и липидами, рассмотрены также в гл. 25.)

4.1. Полисахариды

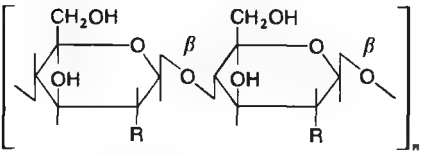
ПЕРВИЧНАЯ СТРУКТУРА

Полисахариды — это полимеры сахаров. К наиболее простым из них относятся линейные гомополимеры, в которых повторяющейся единицей является моносахарид; при этом все связи между сахарами одинаковы. Примерами таких полисахаридов могут служить амилоза, составная часть крахмала, служащего резервным питательным веществом у растений; целлюлоза, основной компонент древесины и других растительных тканей; хитин, из которого построены панцири насекомых и других членистоногих. На рис. 4.1 приведены структурные формулы этих соединений.

Как целлюлоза, так и амилоза являются гомополимерами глюкозы. Их первичные структуры различаются лишь стереохимией глюкозы в положении 1. Свободные сахара

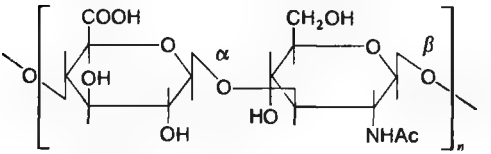


Амилоза

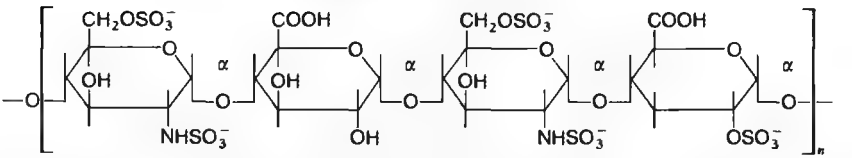


Целлюлоза, R = OH

Хитин, $R = NH-\overset{\overset{O}{\|}}{C}-CH_3$



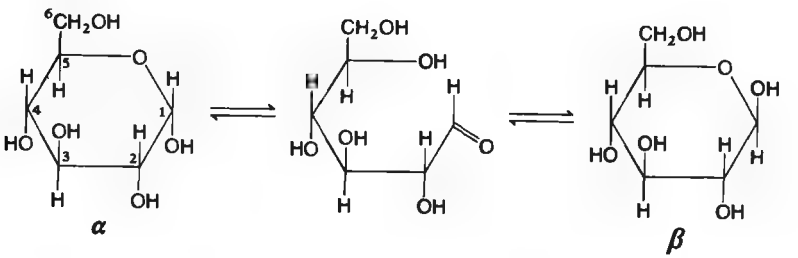
Гиалуроновая кислота



Гепарин

РИС.4.1. Первичная структура нескольких типичных полисахаридных цепей. Система нумерации для гексоз показана на примере амилозы. Заметим, что в том случае, когда моносахарид исходно находится в такой аномерной конфигурации, при которой связи между С-1 и кислородом и между С-6 и кислородом расположены по разные стороиы плоскости сахаридного кольца, образуются α-связи, в противном случае образуются β-связи.

существуют в виде двух аномерных форм, находящихся в равновесии:



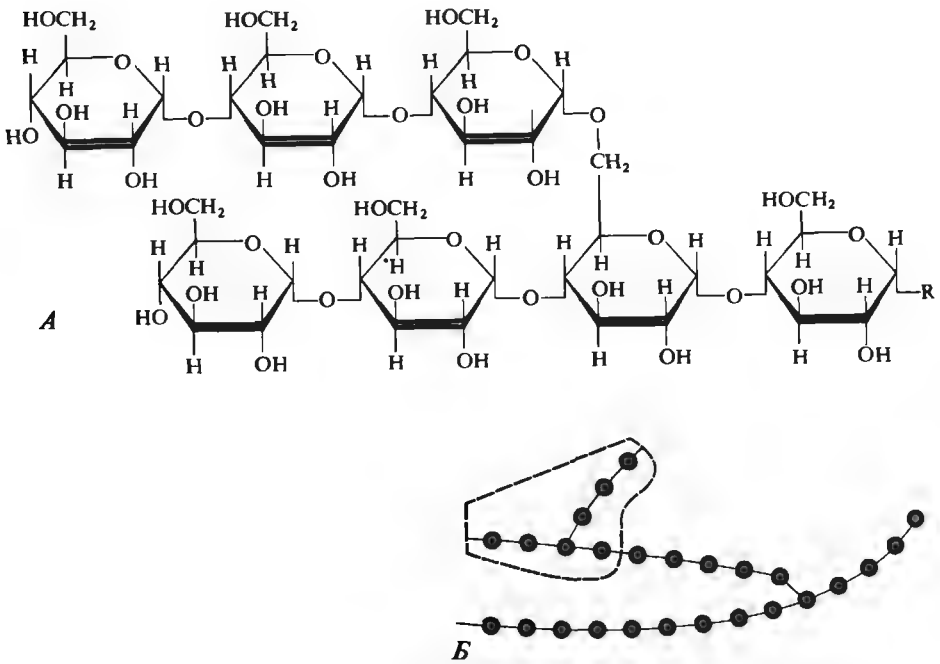


РИС.4.2. Структура гликогена. *А*. Ковалентная структура области ветвления. *Б*. Схематическое изображение большей части полимерной молекулы гликогена; на рис. *А* более детально показана структура участка, обведенного пунктирной линией.

В полисахаридах полуацеталь в положении 1 переходит в значительно менее подвижный ацеталь, в результате чего конфигурация в этом положении становится фиксированной. Мономеры глюкозы в амилозе соединены α-1,4-связями, в то время как для целлюлозы характерны связи β-1,4.

В других линейных полисахаридах основной повторяющейся единицей служит дисахарид, например гиалуроновая кислота (рис. 4.1), в которой чередующиеся N-ацетилглюкозамин и глюкуроновая кислота соединены β-1,4- и α-1,3-связями. Такие чередующиеся полимеры обычно образуют гели. Часто они составляют матрикс, в который погружены волокнистые компоненты клеточных стенок или других опорных клеточных структур. В более сложных полисахаридах последовательность остатков нерегулярна. Например, альгиновые кислоты бурых водорослей могут состоять из чередующихся блоков различных гомополисахаридов. Другие компоненты матриксов сульфонированы в различных положениях. Так, мономерной единицей гепарина (рис. 4.1) является сульфонированный тетрасахарид.

Некоторые полисахариды имеют более сложную первичную структуру из-за наличия ветвлений. В этом отношении они совсем не похожи на белки и нуклеиновые кислоты, молекулы которых всегда линейны или замкнуты в кольцо. Гликоген, резервная форма глюкозы, состоит в основном из линейных α-1,4-полимеризованных участков. Однако эти участки случайным образом связаны между собой α-1,6-ветвями, в результате чего образуется сложная сеть (рис. 4.2). Некоторые очень сложные полисахариды являются составной частью клеточных стенок высших растений. Сюда относятся пектины и гемицеллюлозы. Пектины — это кислые полисахариды, содержащие галактуроновую кислоту и

рамнозу; они различаются по степени разветвленности и степени этерификации карбоксильных групп галактуроновой кислоты. Гемицеллюлозы также являются сложными полисахаридами. Достигнуты некоторые успехи в понимании того, как все эти полимеры организованы в клеточной стенке растений.

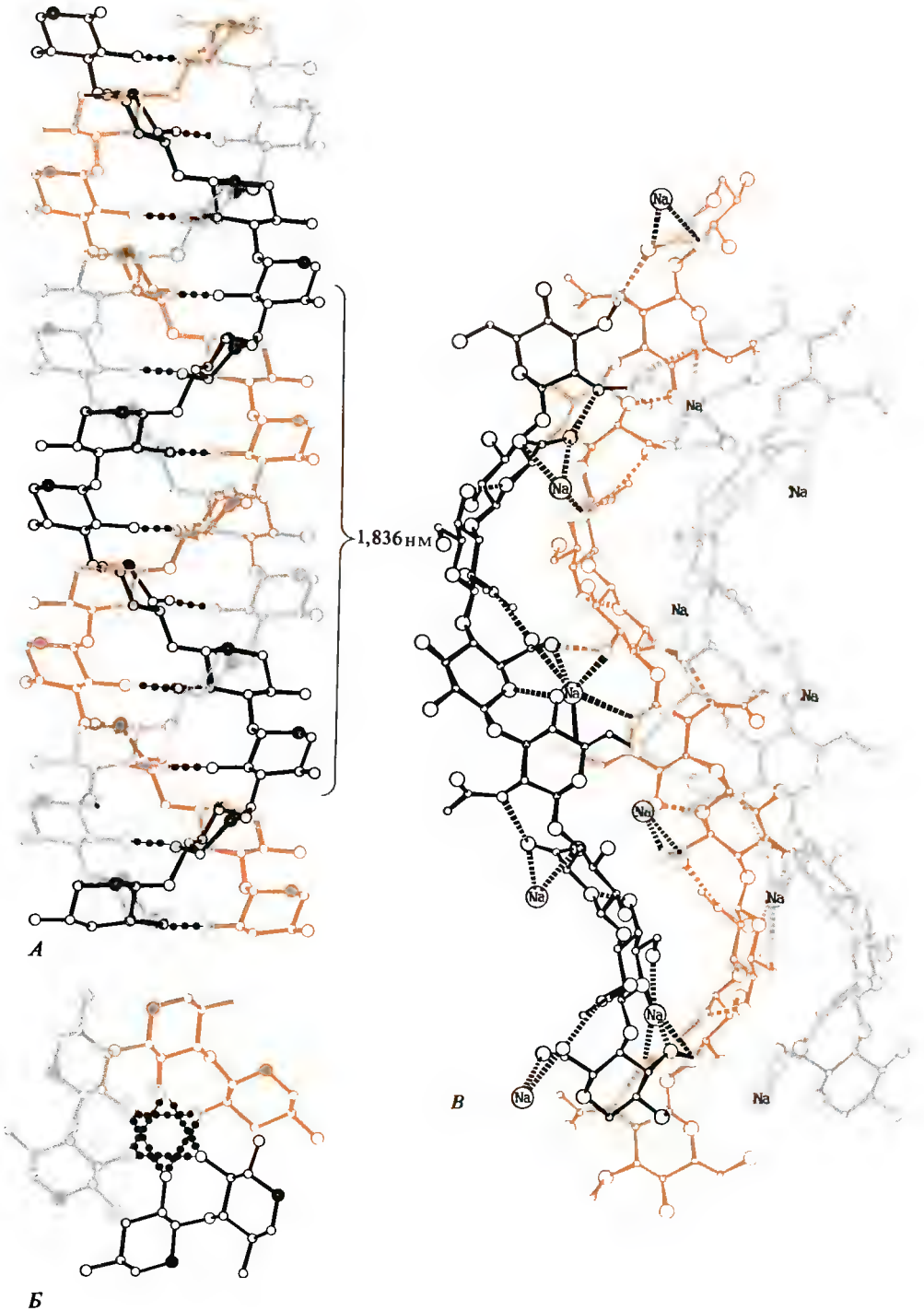
УРОВНИ СТРУКТУРНОЙ ОРГАНИЗАЦИИ ПОЛИСАХАРИДОВ

Конформации полисахаридов весьма разнообразны, поэтому ясно, что понятие о различных уровнях структурной организации, которым мы пользовались при описании строения белков или нуклеиновых кислот, полезно и в данном случае. Как правило, первичная структура полисахаридов весьма регулярна — полимер состоит из больших блоков, в которых повторяются один и тот же остаток или одинаковая последовательность остатков. В этом отношении полисахариды напоминают такой структурный белок, как коллаген.

Волокнистые полисахариды, например хитин и целлюлоза, имеют очень жесткую вторичную структуру. Стерические препятствия, возникающие при взаимодействии громоздких сахарных групп, сильно ограничивают число допустимых конформаций. Волокна в достаточной степени упорядочены, что позволяет получить детальную рентгеновскую дифракционную картину, но лишь некоторые из рентгенограмм удалось расшифровать и построить однозначные структурные модели. Некоторые полисахаридные волокна составлены из лент, содержащих две цепи, которые стабилизированы водородными связями между остатками. Другой тип вторичной структуры полисахаридов представляют спирали. Например, в клеточных стенках некоторых морских водорослей вместо целлюлозы или хитина содержится ксилан (β-1,3-полимер ксилозы). Он образует трехцепочечную спираль, у которой все три цепи параллельны (рис. 4.3, *А*). На один виток спирали приходится шесть остатков, структура стабилизирована центральным ядром межцепочечных водородных связей между гидроксильными группами сахаров (рис. 4.3,*Б*). Таким образом, в общих чертах организация ксилана такая же, как у коллагена.

Структура полисахаридов, образующих матрикс, существенно более лабильна. Каррагенаны представляют собой двойные спирали, при этом на один виток каждой цепи приходятся три дисахаридных остатка. Молекула гиалуроновой кислоты образует левую одиночную спираль с четырьмя дисахаридными остатками на один виток. На рис. 4.3,*В* одна цепь изображена в цвете. Структура стабилизирована водородными связями, образующимися между цепочками. В волокне гиалуроновой кислоты соседние спирали связаны между собой солевыми мостиками и водородными связями. Нарушения регулярности в такой структуре, вызванные либо встраиванием каких-либо других сахаридов, либо ветвлением, существенно влияют на всю конформацию полисахарида. По-видимому, именно эти нарушения обусловливают механическую податливость волокон или гелей, которые образует гиалуроновая кислота.

РИС.4.3. Трехмерная структура двух полисахаридных спиралей, определенная с помощью рентгеновской дифракции от волокон. *А*. β-1,3-ксилан из синезеленых водорослей. Это трехцепочечная спираль; одна цепь изображена в цвете. *Б*. Схематическое изображение водородных связей, соединяющих три цепи между собой. [Atkins et al., Proc. Roy. Soc. Ser. B, **173**, 209 (1969).].
В. Натриевая соль гиалуроновой кислоты. Это одвоцепочечная спиральная структура, у которой на один виток приходятся четыре дисахарида; показано чуть больше одного витка спирали. В волокне образуются межцепочечные и внутрицепочечные водородные связи. Присутствует большое количество координационно связанных ионов Na^+; для некоторых из них показаны связи с кислородом. Для большей наглядности одна цепь полимера изображена в цвете. [Более детально см.: Guss et al., J. Mol. Biol., **95**, 359 (1975).] Рис. *А* и *Б* выполнены по рисункам, предоставленным д-ром Эткинсом. *В* выполнен по рисунку, предоставленному д-ром Уинтером.


1,836 нм
Na
А
Б
В

. Под третичной структурой полисахаридов понимают характер укладки спиралей. Четвертичная структура образуется в результате объединения нескольких полисахаридных молекул. Например, целлюлоза — это, по-видимому, двухцепочечная лента. Последняя может быть свернута в устойчивую регулярную структуру. Упорядоченный ансамбль из свернутых целлюлозных лент, охватывающий иногда всю поверхность растительной клетки, образует четвертичную структуру. Для других полисахаридов характерны желеобразные структуры (гели). Гели, образованные каррагенанами, представляют собой сеть из полисахаридных цепей, соединенных во многих местах двухцепочечными мостиками.

О СТРУКТУРЕ И ФУНКЦИЯХ ПОЛИСАХАРИДОВ

Перечислим некоторые вопросы,касающиеся полисахаридов, и остановимся на трудностях, возникающих на пути изучения этих молекул. Как отдельные модификации в основной ковалентной структуре влияют на вторичную и третичную структуру? Не распределены ли эти модификации каким-либо регулярным образом, приводя к некой упорядоченности, до сих пор ускользавшей от нашего внимания? Каковы особенности синтеза полисахаридов в клетке, приводящие к образованию нужных макроскопических форм и механических свойств? Например, вводятся ли отдельные химические модификации случайно (так, что сохраняется только число модификаций каждого типа) или их пространственное положение строго задано? Какие специфические взаимодействия происходят между различными полисахаридными молекулами и между полисахаридами и другими веществами, находящимися в клетке? Является ли клеточная стенка лишь крупным ассоциатом с множеством случайных свойств, содержащим множество включенных молекул и дефектов, или ей свойственны элементы ближнего или дальнего порядка? Как, исходя из химической структуры полисахаридов, на чисто физическом уровне количественно описать механические свойства полисахаридных материалов?

Все эти вопросы очень важны, но для того, чтобы ответить на большинство из них, надо затратить немало труда. Пока не создано биологических тестов, которые позволяли бы установить структурную целостность полисахаридов, аналогичных тестам, использующимся в случае белков или нуклеиновых кислот. В большинстве проб полисахариды оказываются гетерогенными по длине, а нередко также и по составу. Это серьезно усложняет применение многих физических методов, которые были созданы для работы с гомогенными системами. Для спектроскопических методов, с помощью которых в основном изучают локальную структуру, отсутствие гомогенности не столь существенно, но особенности полисахаридных остатков таковы, что, вообще говоря, большинство спектроскопических методов оказываются непригодными для исследования полисахаридов. Мощным инструментом, с помощью которого сложные системы белков или нуклеиновых кислот можно сделать более удобными для физических исследований, являются генетические методы.

В случае полисахаридов, однако, этот подход оказывается не столь полезным, так как влияние генов на эти вещества не является прямым. В настоящее время, по всей видимости, к наиболее плодотворным подходам следует отнести рентгеноструктурные и электронно-микроскопические исследования полисахаридов интактных клеток (или по крайней мере клеток, как можно менее поврежденных). Необходимым условием для применения большинства физических методов исследования полисахаридов в растворе, обладающих высоким разрешением, является усовершенствование методики получения и очистки соответствующих препаратов.

4.2. Ассоциаты, образованные из разных макромолекул

КОМПЛЕКСЫ САХАРИДОВ С ПЕПТИДАМИ, БЕЛКАМИ ИЛИ ЛИПИДАМИ В КЛЕТОЧНЫХ СТЕНКАХ БАКТЕРИЙ

Обнаружено, что ряд сложных полимерных структур в стенках бактерий содержит олигосахариды или полисахариды в комплексе с другими веществами. У всех известных бактерий, за исключением микоплазм, в качестве структуры, поддерживающей их форму, выступает пептидогликановая сеть. Материал этой сети, называемый также муреином, гликопептидом или мукопептидом, содержит длинные цепи полисахарида, у которого повторяющейся единицей служит β-1,4-N-ацетилглюкозамин—β-1,4-N-ацетилмурамовая кислота (рис. 4.4,*А*). Эти цепи связаны друг с другом короткими олигопептидными мостиками. Последовательность пептидов и характер поперечного связывания зависят от вида бактерий.

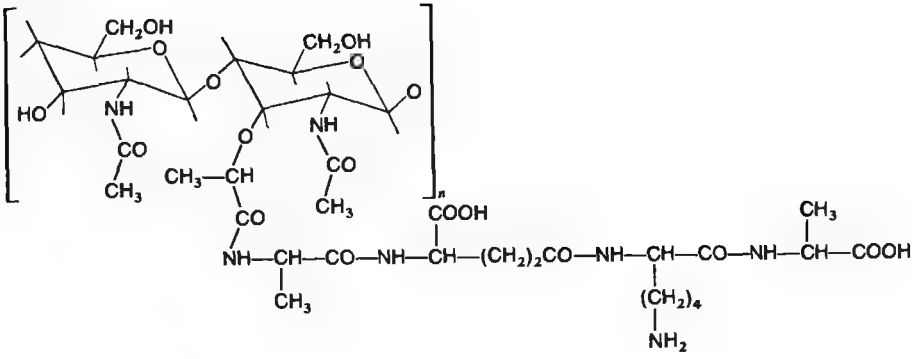
Бактерии с клеточными стенками разного типа дают разную реакцию по Граму (обработка кристаллическим фиолетовым и иодом, а затем отмывка в органическом растворителе). Грамотрицательные бактерии при такой обработке полностью обесцвечиваются. Их оболочка состоит из наружной стенки, которая представляет собой типичный липидный бислой толщиной 80 Å, из пептидогликанового слоя толщиной 20-30 Å (их разделяет промежуток в 75 Å — периплазматическое пространство) и, наконец, из внутренней стенки — еще одного липидного бислоя толщиной 80 Å. Пептидогликановый слой типичной грамотрицательной бактерии, например *Escherichia coli* (рис. 4.4,*Б*), имеет небольшое число поперечных сшивок. Он содержит около $3 \cdot 10^6$ повторяющихся единиц и составляет от 5 до 10% веса сухого вещества клеточной стенки. С наружной его поверхностью связаны липопротеидные молекулы, которые соединяют пептидогликан с наружной стенкой. Эти молекулы присоединены своими С-концами к аминокислотам пептидогликана, а иа их модифицированном N-конце находятся три жирнокислотные цепи.

Грамположительные бактерии после обработки по Граму остаются окрашенными в насыщенный темно-синий цвет. У них есть наружная клеточная стенка толщиной от 100 до 500 Å и единственная липидная бислойная внутренняя стенка толщиной 80 Å. Наружная стенка — это в основном пептидогликан, который составляет от 40 до 90% веса ее сухого вещества. Этот пептидогликан содержит множество поперечных сшивок, как показано на примере *Staphylococcus aureus* (рис. 4.4,*Б*).

У многих грамположительных бактерий с пептидогликаном связаны тейхоевые кислоты — вещества типа полиглицерофосфата. Иногда они бывают соединены фосфодиэфирными мостиками с 6-гидроксильными группами остатков мурамовой кислоты. Одна из детально исследованных бактериальных структур такого рода принадлежит *Streptococcus faecalis:* это липотейхоевая кислота, содержащая липополисахарид, связанный с тейхоевой кислотой (рис. 4.5). По-видимому, остатки тейхоевой кислоты не играют никакой роли в придании жесткости клеточной стенке; возможно, они участвуют в связывании и переносе двухвалентных катионов.

У *S. faecalis* с тейхоевой кислотой связан очень простой липополисахарид; в наружных стенках грамотрицательных бактерий присутствуют существенно более сложные полисахариды с мол. массой свыше 10 000. Согласно приближенной модели, липополисахарид *Salmonella typhimurium* — это сложная разветвленная структура, содержащая множество сахаридов и жирных кислот. О конформации или о других физических свойствах такой структуры почти ничего неизвестно. Возможно, даже эта модель слишком упрощена — известно, например, что у *E. coli* с поверхностным липополисахаридом могут быть ковалентно связаны белки.

S. aureus



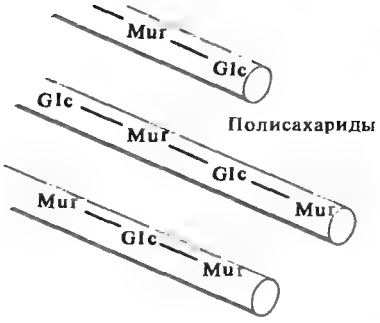
GlcNAc-MurNAc
|
L-Ala-D-isoGln-L-Lys-D-Ala

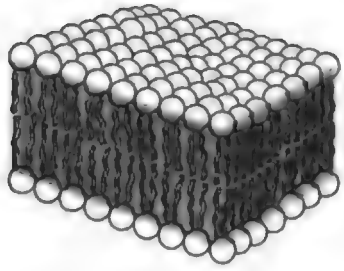
E. Coli

GlcNAc-MurNAc
|
L-Ala-D-isoGlu-Dpm- -Ala

A

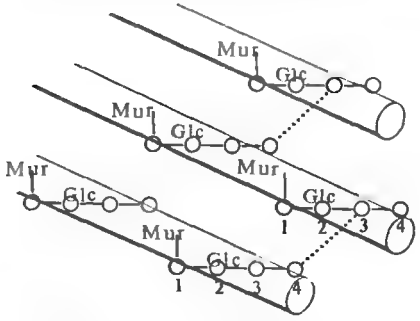
РИС. 4.4. Элементы структуры пептидогликановой клеточной стенки двух бактерий: *Escherichia coli* — типичного грамотрицательного организма, и *Staphylococcus aureus* — типичного грамположительного организма. *А*. Химическая структура пептидогликановой цепи. Она состоит из повторяющихся дисахаридных единиц N-ацетилглюкозамин — N-ацетилмурамовая кислота, причем к последней присоединены короткие пептидные цепи, содержащие как D-, так и L-аминокислоты. Dpm — это (*мезо*)-диаминопимелиновая кислота. *Б*. Схематическое изображение клеточной стенки, которая состоит из слоев сшитых пептидогликанов и липидных бислоев. У *E. coli* с пептидогликанами ковалентно связаны липопротеиды. Схема строения этого белка, изображенная на рисунке, предложена МакЛакланом исходя из аминокислотной последовательности. [A.D. McLachlan, J. Mol. Biol., **122**, 493 (1978).] Структура липотейхоевой кислоты из клеточной стенки *S. aureus* чисто предположительная. [Более подробно см.: Braun, Hantke, 1974, Stanier et al., 1976; DiRienzo et al., 1978.] (Рисунок Ирвинга Гейса.)



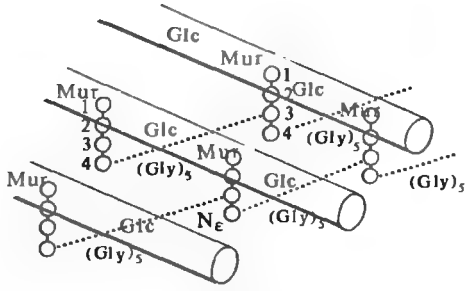


Липидный бислой

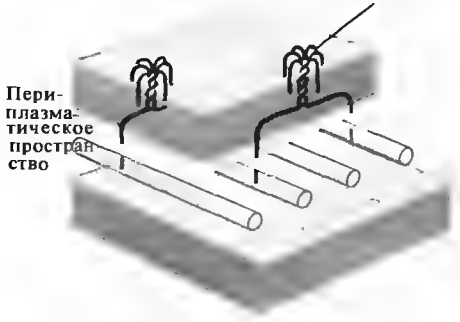
Грамотрицательная бактерия: *E. coli*



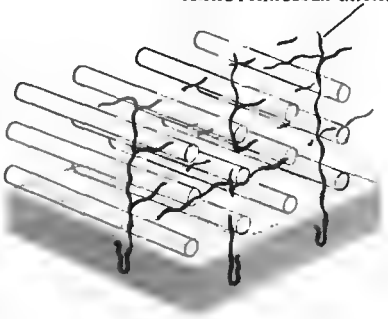
Грамположительная бактерия: *S. aureus*



Липопротеидный димер



Липотейхоевая кислота



Б

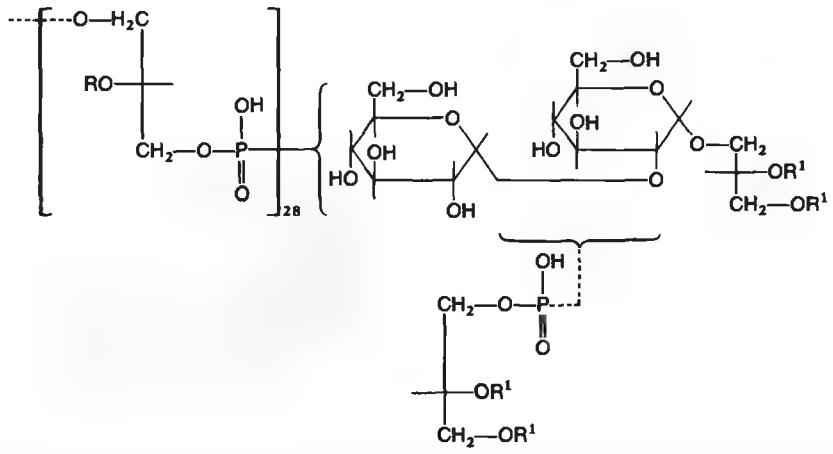


РИС.4.5. Химическая структура липотейхоевой кислоты из *S. faecalis*. R — диглюкозил или атом водорода; R^1 — жирная кислота. Точное место присоединеиия к сахариду изображенного внизу фосфолипида неизвестно. [V. Braun, K. Hantke, Ann. Rev. Biochem., **43**, 89 (1974).]

ГЛИКОПРОТЕИДЫ НА ПОВЕРХНОСТИ ЖИВОТНЫХ КЛЕТОК

Животным клеткам не нужна клеточная стенка ни для придания им жесткости, как клеткам растений, ни для защиты от осмотического шока, как бактериям. Эти клетки ограничены липидной бислойной мембраной, содержащей большое количество белков, лишь малая часть которых связана с олигосахаридами, образуя гликопротеидный комплекс. Содержание сахара может варьировать от нескольких остатков на большую молекулу белка до величины, сравнимой по массе с белковой частью. Остатки сахара могут быть связаны как с одним аминокислотным остатком, так и с несколькими. Так, к примеру, бычья рибонуклеаза B отличается от рибонуклеазы A (гл. 2) тем, что она содержит олигосахарид

$$(\text{Манноза})_6\text{—}(\text{N-ацетилглюкозамин})_2\text{ —}$$

специфически связанный с Asn^{34}. Хорионный гонадотропин человека содержит дисахарид

$$\text{Галактоза-}\beta\text{ (1-3)-N-ацетилглюкозамин—}$$

связанный с остатками серина 116, 121 и 123.

У гликофорина (основного гликопротеида мембраны эритроцита) сахариды присоединеиы к одному остатку аспарагииа, семи остаткам сериив и восьми остаткам треонина (рис. 2.16), причем каждый их этих сахаридов имеет сложную структуру (рис. 4.6). На долю всех этих сахаров приходится примерно 60% массы гликофорина. Еще один гликопротеид с высоким содержанием сахара выделен из антарктических рыб. У этого белка есть замечательное свойство — он понижает температуру замерзания воды значительно сильнее, чем можно было бы ожидать, исходя из обычных коллигативных свойств растворов. Единицей, повторяющейся 31 раз, является трипептидная последовательность Thr—Ala—Ala. К каждому остатку треонина присоединен дисахарид галактоза-β(1-3)-N-ацетилгалактозамин.

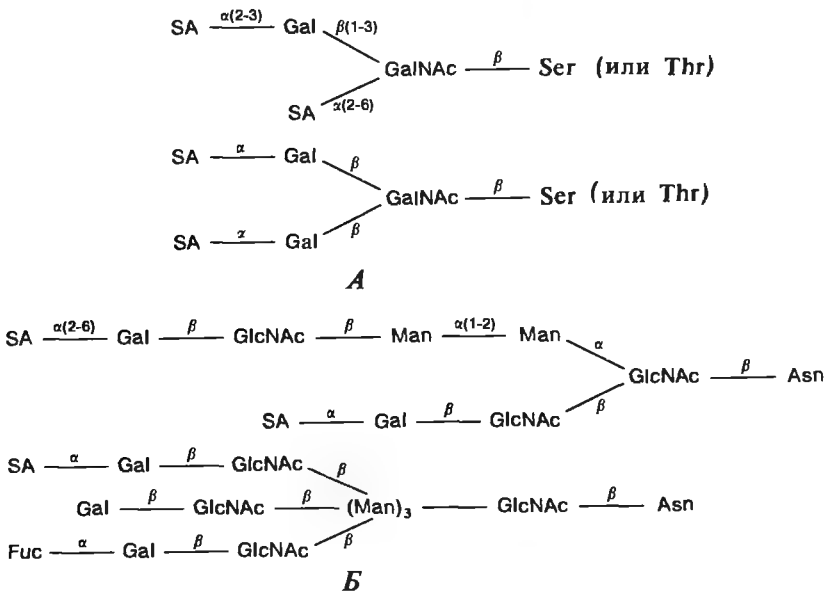


РИС. 4.6. Возможная структура сахаридных частей гликофорина из эритроцита человека. (Относительно предполагаемых мест присоединения этих сахаров см. рис. 2.16.) *А*. Две альтернативные структуриые модели сахаридов, присоединяющихся к серину или треонину. *Б*. Две альтернативные структурные модели сахаридов, присоединяющихся к аспарагину. Gal — галактоза; SA — сиаловая кислота; GlcNAc — N-ацетилглюкозамин; GalNAc — N-ацетилгалактозамин; Fuc — фукоза; Man — манноза. [V.T. Marchesi et al., Ann.Rev.Biochem., **45**, 667 (1976).]

Структура сахаридов очень разнообразна. Представление о некоторых довольно распространенных, но отнюдь не универсальных особенностях строения этих соединений дает рис. 4.7. Сахариды, связанные с серином или треонином, часто содержат дисахарид

Галактоза-β(1-3)-N-ацетилгалактозамин

а олигосахариды, связанные с аспарагином, — трисахарид

Манноза-β(1-4)-N-ацетилглюкозамин-β(1-4)-N-ацетилглюкозамин.

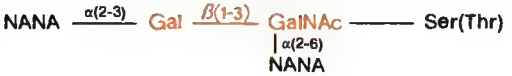
Олигосахарид, связанный с аспарагином, называется простым, если он содержит только эти два типа сахаров, и сложным, если присутствуют другие типы мономеров, такие, как галактоза, фукоза и сиаловая кислота.

Влияние сахаридиых остатков на свойства связанных с ними белков практически неизвестно. По крайней мере некоторые гликопротеиды очень устойчивы к тепловой денатурации. Поведение большинства белков, связанных с большим количеством сахаров, оказывается аномальным при исследовании их теми методами, которые применяются для определения размера и формы молекул.

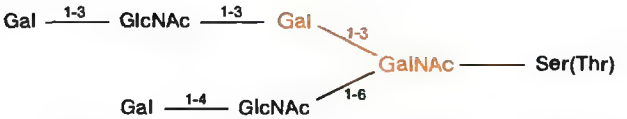
Не совсем ясно и то, какую биологическую функцию выполняют сахаридные части гликопротеидных молекул. Сахара — это сильные антигены, которые как бы ставят индивидуальную метку на клетки данного организма, с тем чтобы защитные иммунные механизмы могли отличить эти клетки от чужеродных. Сахара, вероятно, участвуют в меж-

Гликопротеид-антифриз, человеческий IgA_1

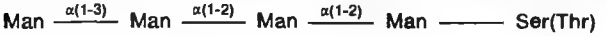
Некоторые муцины, секретируемые подчелюстной железой



Фетуин



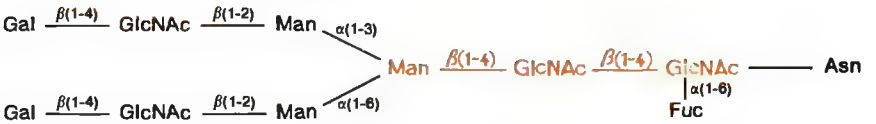
Кор группоспецифического вещества крови человека



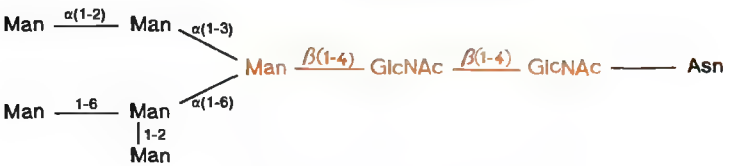
Маннан из *S. cerevisiae*

А

Бычья РНКаза В

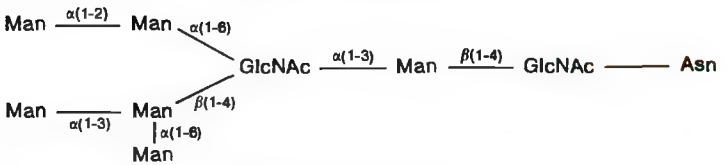


Бычий IgG



α-Амилаза *A. oryza*, куриный овальбумин

Бромелаин ананаса



Человеческий IgE

Б

клеточном узнавании, являясь компонентами рецепторов клеточной поверхности для белков, элементов иммунной системы или даже для молекул меньшего размера. Жесткость сахаридных структур, их растворимость в воде и огромное разнообразие возможных конфигураций, очевидно, делают их идеальными для выполнения этих функций. Не все гликопротеиды находятся на клеточной поверхности. Многие из них обнаруживаются в сыворотке, выделяются в пищеварительный тракт и т.д. Можно предположить, что сахарид — это нечто вроде билета, который разрешает перенос белка или препятствует его разрушению клетками до тех пор, пока сахарид остается связанным с белковой молекулой.

НУКЛЕОПРОТЕИДЫ: НЕКОВАЛЕНТНЫЕ КОМПЛЕКСЫ МЕЖДУ БЕЛКАМИ И НУКЛЕИНОВЫМИ КИСЛОТАМИ

Комплексы, образуемые белками и нуклеиновыми кислотами, находятся в центре тех процессов, в ходе которых в клетке воспроизводится и используется генетическая информация. Взаимодействия между белками и нуклеиновыми кислотами принимают множество форм — от простых случаев, в которых РНК или ДНК выступают в качестве субстратов ферментов, до взаимодействий в сложных упорядоченных ансамблях между белками и одной или несколькими нуклеиновыми кислотами. В числе последних — три типа систем, которые изучаются особенно широко: вирусы, хроматин и рибосомы (табл. 4.1).

Таблица 4.1

ХАРАКТЕРИСТИКИ НЕКОТОРЫХ НУКЛЕОПРОТЕИДНЫХ СИСТЕМ

	Нуклеосома (хроматин)	70 S-рибосома *E. coli*	Вирус табачной мозаики	Фаг P22	Аденовирус
Мол. масса	$2{,}4 \cdot 10^5$	$2{,}6 \cdot 10^6$	$4{,}0 \cdot 10^7$	$\sim 5{,}5 \cdot 10^7$	$\sim 1{,}6 \cdot 10^8$
Содержание белка, % (по весу)	45	33	95	50	87
Число молекул и тип нуклеиновых кислот	1 ДНК	3 РНК	1 РНК	1 ДНК	1 ДНК
Число белков	9	57	2130	~500	>2500
Число разных белков	5	53	1	6	~12

Большинство белков, присутствующих в нуклеопротеидных комплексах, достаточно малы: их мол. масса составляет от 10 000 до 25 000. Общей чертой таких комплексов является то, что входящие в их состав белки связываются с единственной молекулой нуклеиновой кислоты и, возможно, даже с одним и тем же ее участком. По преимуществу это основные белки. В этом есть смысл, так как их положительно заряженные группы могут частично нейтрализовать полианион, каким является нуклеиновая кислота, и образовать с ней стабильный комплекс. Однако не все белки, прочно связывающиеся с нуклеиновыми кислотами, являются основными.

РИС. 4.7. Типичные сахариды, входящие в состав гликопротеидов. Структурные элемеиты, общие для многих белков, показаны в цвете. *А*. Сахариды, присоединяющиеся к треонину или серину. *Б*. Сахариды, присоединяющиеся к аспарагину. NANA — N-ацетилнейраминовая кислота; Xyl — ксилоза; другие сокращения объяснены в подписи к рис. 4.6. [R. Kornfeld, S. Kornfeld, Ann.Rev.Biochem., 45, 217 (1976).]

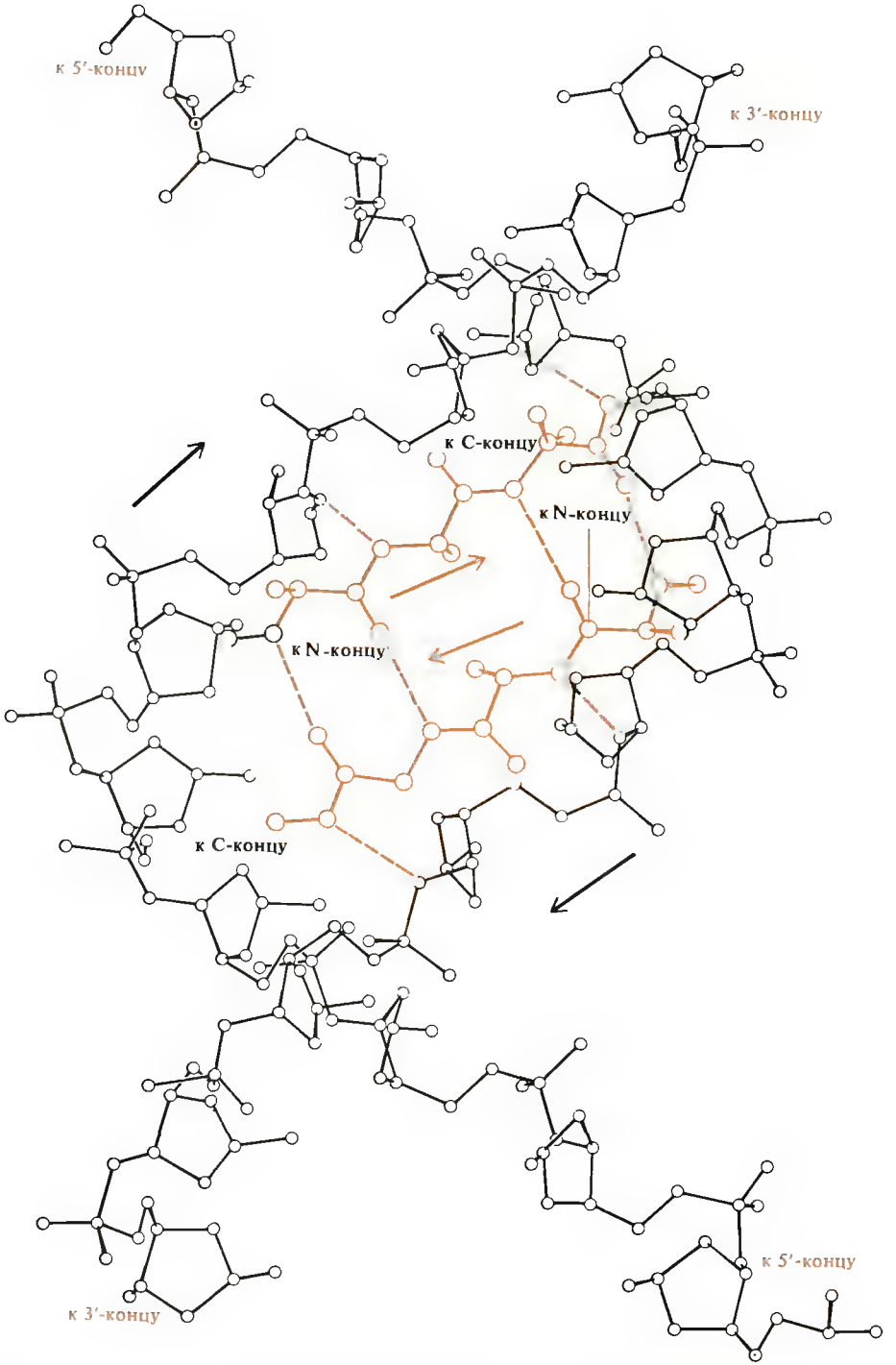
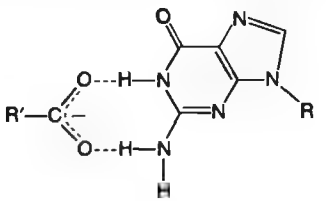


РИС. 4.8. Возможная модель взаимодействия между антипараллельным складчатым слоем и двухцепочечной ДНК. Слой помещается в малом желобке ДНК; оси симметрии второго порядка складчатого слоя и лежащей под ним ДНК совпадают. Водородные связи образуются между NH-группами пептидов и O-3′-атомами дезоксирибозы. [G.M. Church et al., Proc. Nat. Acad. Sci. USA; 74, 1458 (1977).]

При исследовании нуклеопротеидных комплексов возникают такие же вопросы, что и в случае других структур, образованных из разных макромолекул. Так, хотелось бы знать долю контактов белок—белок и нуклеиновая кислота—нуклеиновая кислота в сравнении с долей гетерогенных контактов. Предельными случаями могли бы быть такие контакты между молекулой нуклеиновой кислоты и мультипротеидным ядром или между белковыми молекулами и полинуклеотидным ядром.

Какие существуют специфические взаимодействия между нуклеотидами и аминокислотными остатками? Если здесь и есть какой-либо простой код, с помощью которого эти молекулы узнают друг друга, то пока его не удалось расшифровать. Весьма распространены в таких системах тесное взаимодействие между лизином либо аргинином и фосфатами и интеркаляция плоских ароматических аминокислотных остатков между основаниями полинуклеотидной цепи. Использование молекулярных моделей показывает, что форма β-слоев позволяет им хорошо разместиться в одном из желобков двухцепочечной ДНК (рис. 4.8). Согласно другим модельным построениям, α-спирали, содержащие в большом количестве аргинин или лизин, могут располагаться в большом или малом желобках двойной спирали ДНК и нейтрализовать фосфаты на обеих сторонах желобка.

Образование водородных связей между аминокислотами и основаниями может приводить к специфическим, зависящим от последовательности взаимодействиям между белками и одноцепочечными нуклеиновыми кислотами. Например, в смесях ДМСО — вода карбоксилат-анионы (моделирующие боковые группы аспарагина и глутамина) образуют комплекс с гуанином, в 30 раз более прочный, чем комплекс между гуанином и цитозином в тех же условиях. Среди пяти обычных оснований нуклеиновых кислот только гуанин может вступать в это взаимодействие.



Такое взаимодействие объясняет, почему кислые белки образуют прочные комплексы с нуклеиновыми кислотами.

Изменяются ли физические и химические свойства белков или нуклеиновых кислот при их объединении в нуклеопротеидный ансамбль? Как правило, на этот вопрос можно ответить утвердительно, но во многих случаях неизвестно, происходят ли такие изменения благодаря просто конформационным перестройкам, вызванным одним компонентом в другом, или имеют место какие-то более прямые эффекты. К примеру, проявляются ли какие-либо особые свойства аминокислот или нуклеотидов при их непосредственном контакте? Большинство нуклеиновых кислот существуют только в виде комплексов с белками, последние же могут находиться и в свободном состоянии. Интересно выяснить, обладают ли белки, связывающиеся с нуклеиновыми кислотами, какими-либо общими свойствами,отличными от свойств других растворимых белков. Есть указания на то, что многие белки, которые связаны с нуклеиновыми кислотами, менее глобулярны, чем обычные небольшие белки, но пока неясно, насколько универсально это свойство.

БЕЛКИ И НУКЛЕИНОВЫЕ КИСЛОТЫ ВИРУСОВ

Обычно вирусы состоят из нуклеиновой кислоты и окружающей ее белковой оболочки. Простейший палочковидный вирус, состоящий из единственной одноцепочечной РНК и множества копий одного белка, — вирус табачной мозаики. Белок без участия РНК соби-

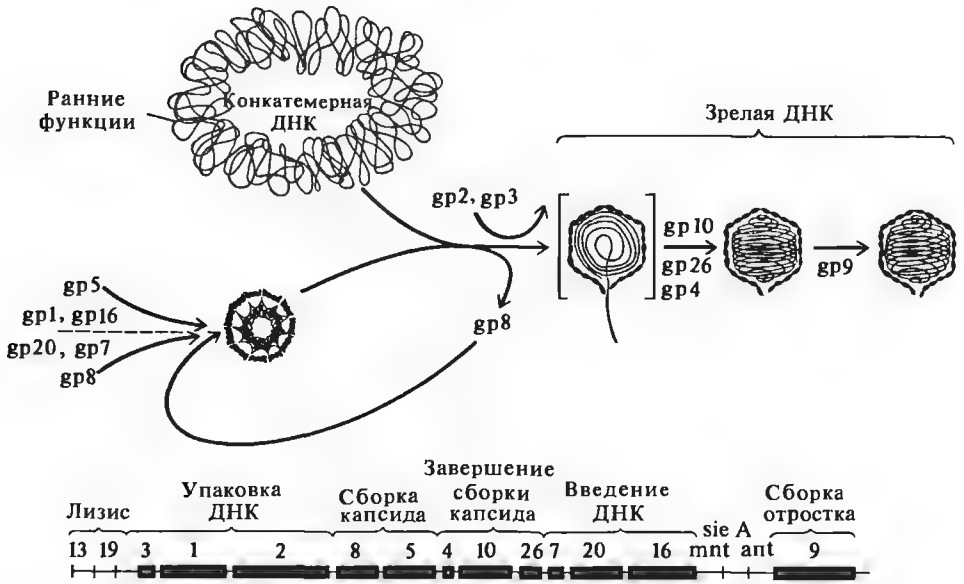


РИС. 4.9. Схема сборки бактериофага P22, инфицирующего *S. typhimurium*. gp—продукты генов фага. На первом этапе образуется сферический прокапсид, состоящий в основном из gp5 и gp8. Когда в капсид проникает часть конкатемерной ДНК, gp8 (белок, ответственный за поддержание формы) уходит. Внутри капсида ДНК укорачивается до нужной длины. Затем осуществляется сборка отростка фага. [J. King et al., Cell, **15**, 551 (1978).]

рается в структуру, фактически идентичную природному вириону (гл. 2). Таким образом, хотя белок и находится снаружи, по-видимому, именно он является организующим ядром структуры, а РНК лишь определяет ее длину. РНК у этого вируса сильно вытянута. Стэкинг между основаниями отсутствует, зато основания взаимодействуют с белком. Таким образом, во время сборки вируса структура РНК должна коренным образом измениться по сравнению с той, что была в растворе.

Многие вирусы бактерий (бактериофаги) состоят из белковой головки, или капсида, в которую заключена ДНК или РНК. На рис. 4.9 показаны этапы сборки одного из таких вирусов. Иногда капсид собирается независимо, а затем в него вводится ДНК практически без изменения свойств капсида. Известны и другие случаи, когда введение ДНК сопровождается заметными изменениями в строении и составе капсида. И наконец, есть случаи, когда размер капсида, по-видимому, определяется размером находящейся в нем нуклеиновой кислоты. В более сложных вирусах помимо белковой оболочки имеется и белковое ядро.

Видимо, природа использует целый набор механизмов, чтобы организовать сборку вирусных нуклеопротеидов. Для всех структур характерно наличие множества контактов белок — белок; об этом можно судить не только по внешнему виду вирусных частиц, но и по тому, что их сборка может осуществляться в отсутствие нуклеиновых кислот. За исключением палочковидных вирусов, множество контактов имеется и на уровне нуклеиновой кислоты; в них, вероятно, принимают участие некоторые внутренние белки. Таким образом, в этих вирусах белковые и нуклеиновые компоненты в значительной степени разделены в соответствии с тем, что белок выполняет функцию упаковки для переноса, защиты и введения в клетку нуклеиновой кислоты. Некоторые ферменты, участвующие в сборке ви-

русов, чувствительны к состоянию организации вирусной частицы. К примеру, белки, которые фрагментируют длинные молекулы ДНК-предшественников, не функционируют в случае изолированной нуклеиновой кислоты — для них требуется присутствие заполненного капсида.

ВЗАИМОДЕЙСТВИЕ МЕЖДУ РНК И БЕЛКАМИ В РИБОСОМАХ

Общая картина белково-нуклеиновой организации рибосом сильно отличается от таковой для вирусов. Для нее характерно наличие множества контактов между белками и нуклеиновыми кислотами, причем эти контакты, по-видимому, являются очень тесными. Например, многие рибосомные белки связываются с определенными участками рибосомных РНК (рРНК), защищая их от ферментативного расщепления. Имеют место и коопе-

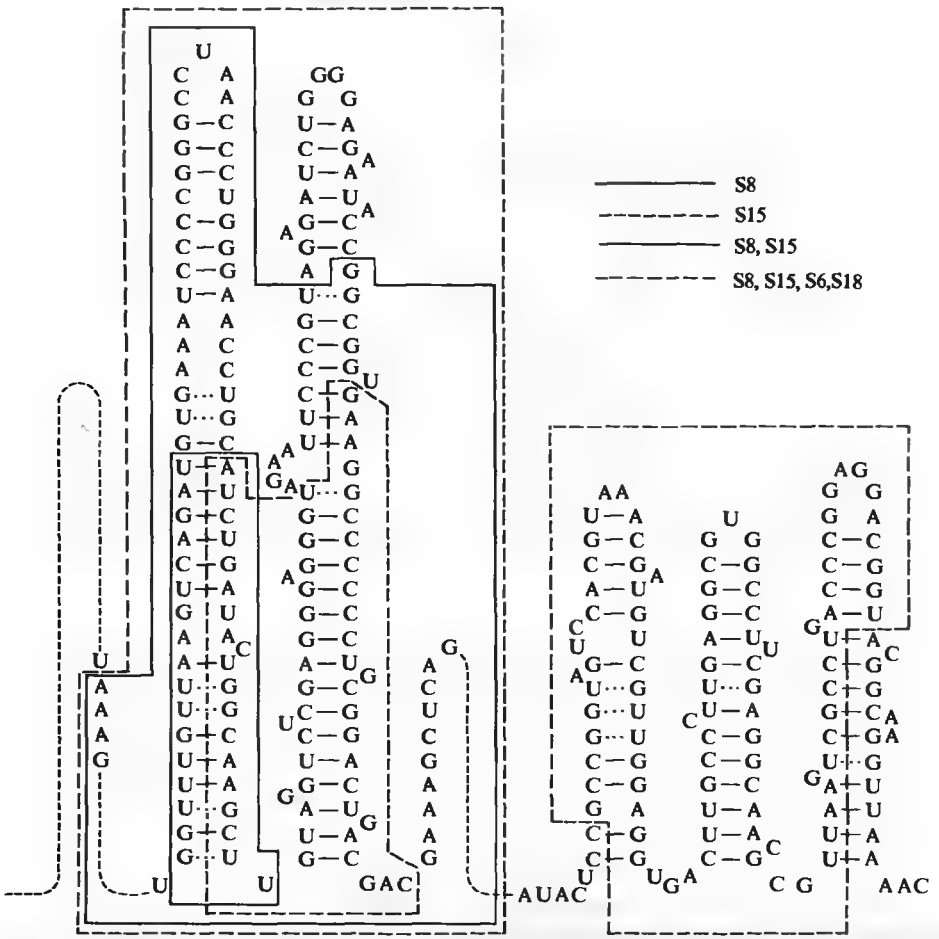


РИС. 4.10. Возможная вторичная структура вебольшой части 16S-рРНК *E. coli*. Указаны области, защищенные от расщепления нуклеазой, в тех случаях, когда к РНК добавлены по отдельности белки S8 и S15 (1), когда эти белки добавлены одновремеино (2) и когда одновременно присутствуют белки S8, S15, S6 и S18 (3). [R. Zimmermann et al., Nucleic Acids Research, 2, 279 (1975).]

ративные эффекты, когда с нуклеиновой кислотой одновременно связывается более одной молекулы белка (рис. 4.10).

Согласно современным представлениям, рибосомные белки в отсутствие рРНК вряд ли могут самостоятельно собираться в какую-либо упорядоченную макроструктуру. Действительно, в растворе удалось обнаружить весьма немного специфических парных контактов между рибосомными белками. В опытах по сшиванию, проведенных на нативных рибосомах, показано, что многие белки расположены достаточно близко друг к другу, однако неизвестно, сколько белковых пар непосредственно контактирует и являются ли эти контакты достаточно прочными, чтобы внести существенный вклад в стабилизацию конечной структуры. (На рис. 1.7 приведена схематическая картина белково-нуклеиновой организации 70S-рибосомы и различных 50S-частиц.)

Свободные рибосомные РНК обладают четко выраженной вторичной, а также, вероятно, и третичной структурой. Большинство данных свидетельствует о том, что хотя при упаковке рРНК в рибосоме происходит существенная компактизация молекулы, ее вторичная структура изменяется незначительно. Многие наблюдения наводят на мысль, что по сравнению со свободными РНК число контактов РНК — РНК в рибосоме увеличивается незначительно. Если это так, то основным фактором, стабилизирующим глобулярные рибосомные субъединицы, должно быть взаимодействие между РНК и белками.

Конформация рибосомных субчастиц очень лабильна. Их свойства существенно изменяются при связывании с некоторыми небольшими белками, тРНК или антибиотиками. Мутации, затрагивающие белки одной субчастицы, иногда сказываются на функционировании другой субчастицы. В системах меньшего размера эффекты такого рода обычно не наблюдаются. Если эти эффекты имеют биологический смысл, это может объяснить, почему природа для определенных целей создала такие сложные структуры.

Свойства отдельных рибосомных компонентов в значительной мере зависят от состояния всей системы. Например, лизиновые группы одного из белков в свободном состоянии химически неактивны, но становятся довольно активными в рибосоме. Белок колицин E_3 вызывает специфическое расщепление 16S-рРНК из 30S-субчастицы только в интактной 70S-частице, но не в свободной 30S-субчастице или свободной РНК. Именно на уровне рибосомы или ее субчастиц проявляются способность к связыванию и ферментативная активность. Было бы интересно узнать, выполняются ли эти функции отдельными белками (при условии, что ансамбль переводит их в активную конформацию) или же многочисленные компоненты на самом деле действуют согласованно на некоторых каталитических этапах белкового синтеза.

ХРОМАТИН: ИЕРАРХИЯ ОРГАНИЗОВАННЫХ СТРУКТУР

За последние несколько лет произошли значительные изменения в наших представлениях о структуре хроматина. Существует несколько уровней ее организации. Комплекс ДНК с пятью типами гистонов образует регулярную упорядоченную структуру. Ее основной единицей является нуклеосома, по-видимому, содержащая (1) по две молекулы каждого гистона H2a, H2b, H3 и H4, (2) одну молекулу гистона H1 и (3) около 200 пар оснований двухцепочечной ДНК. Нуклеосомы расположены вдоль молекулы ДНК, как бусины, при этом каждая из них отделена от соседних участком свободной ДНК.

Существующие модели подтверждают представление о нуклеосомной кор-частице, содержащей (1) около 140 пар оснований ДНК и (2) все гистоны, кроме H1. Кор-частица обладает осью симметрии 2-го порядка. Есть убедительные данные, что гистоны находятся в ее центре. Обнаружено, что некоторые пары гистонов образуют в растворе стабильные димерные комплексы, наблюдались также и тетрамерные белковые комплексы. Вероятно, ДНК намотана регулярным образом вокруг этих белковых тетрамеров. Имеющихся

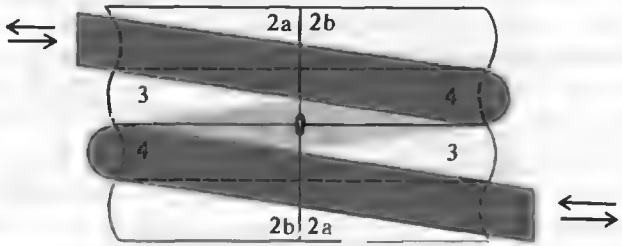


РИС. 4.11. Схема строения повторяющейся субъединицы хроматина эукариот. Изображенная структура (кор-частица) является чисто умозрительной, но весьма правдоподобной. Субъединица содержит по два из четырех упомянутых в тексте гистонов. Указано возможное положение оси симметрии C_2. Такая симметричная организация структуры вполне вероятна, если учесть состав белков и наличие в ДНК оси псевдосимметрии C_2. Данное расположение белков согласуется с тем, что известно о белковых подсистемах и сшивках между белками, но это отнюдь не единственная возможная интерпретация имеющихся данных. Как показано, ДНК имеет полтора тороидальных сверхвитка (см. гл. 24).

данных об организации ДНК достаточно для того, чтобы строить детальные модели, подобные той, что показана на рис. 4.11, но их не хватает, чтобы среди этих моделей выбрать единственную. Кроме того, есть доказательства, что нуклеосомы, содержащие активно транскрибирующиеся гены, имеют более развернутую структуру, чем нуклеосомы с неэкспрессирующимися генами.

В кор-частице нуклеосомы ДНК упакована так, что ее протяженность составляет только 1/7 длины молекулы ДНК в В-форме. Играет ли роль в такой упаковке взаимодействие ДНК—ДНК — неясно. Известно, однако, что в некоторых условиях можно компактизовать ДНК и в отсутствие гистонов.

Природный хроматин помимо гистонов и ДНК содержит в различных количествах множество других белков. Они называются негистоновыми белками как бы в ознаменование того, что мы о них почти ничего не знаем. Их количество и тип зависят от ткани и меняются в широких пределах. Одни из этих белков, видимо, играют роль в экспрессии генов, другие принимают участие в формировании структур более высокого порядка. Влияют ли эти белки на структуру или на положение отдельных нуклеосом — неизвестно.

Существуют четкие указания на то, что in vivo (и в концентрированных гелях in vitro) цепочки нуклеосом сворачиваются в спиральные соленоидоподобные образования. Последние в свою очередь организуются в еще более сложные структуры. В метафазных клетках хроматин упакован в специфические структуры, называемые хромосомами. Детали механизма упаковки и конечная структура хромосом — это важные проблемы, которые предстоит решать следующему поколению биофизических химиков. Необходимо подчеркнуть, что у прокариот, по-видимому, нет хроматина, однако их ДНК тоже уложена в компактные структуры, о которых также далеко не все известно.

4.3. Липиды в биологических мембранах

Все живые клетки окружены наружной мембраной, которую иногда называют плазматической. Эта структура отделяет содержимое клетки от внешней среды. Она действует как своего рода фильтр, пропускающий одни молекулы (а иногда и способствующий их проникновению) и задерживающий другие. Кроме того, на наружной мембране бывают расположены рецепторы, с помощью которых клетка сообщается с внешней средой и с другими клетками.

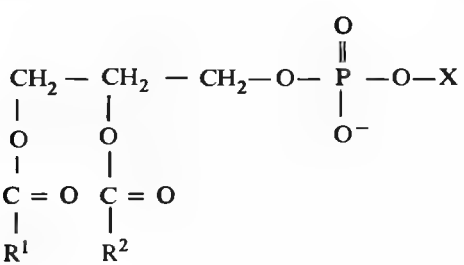
Однако мембранные структуры можно найти не только на периферии клетки. Например, в эукариотических клетках имеется мембрана, которая отделяет ядро от окружающей цитоплазмы. Эукариотические клетки имеют также обширную систему мембранных каналов, называемую эндоплазматическим ретикулумом. Существуют и другие мембраны, которые образуются в особых случаях, например миелиновая оболочка вокруг аксона и так называемые диски в наружном сегменте палочек сетчатки.

Структурную основу мембраны составляют амфифильные липидные молекулы. Амфифильные соединения содержат два основных компонента: полярную группу и более протяженную неполярную часть. Молекулы этого типа легко образуют двойные слои, которые самопроизвольно сворачиваются в мембранные пузырьки (везикулы). В биологических мембранах в этих бислоях располагаются молекулы ряда белков.

ЛИПИДНЫЕ КОМПОНЕНТЫ МЕМБРАН

С точки зрения химического состава отличительной чертой мембран является высокая концентрация в них липидов. Эти молекулы легко экстрагируются из мембран с помощью органических растворителей и плохо растворяются в воде. Мембранные липиды — это в основном особые амфифильные молекулы, имеющие длинные углеводородные хвосты, а также холестерин или его эфиры. Особенности этих молекул и их взаимодействия в значительной степени определяют структуру и свойства биологических мембран.

В большинстве случаев амфифильные соединения — это фосфолипиды (эфиры глицерофосфата), имеющие следующую структуру:



где жирные кислоты (R^1COOH и R^2COOH) этерифицированы по двум гидроксильным группам глицерина, а полярная группа X присоединена к фосфату. Алкильные группы R^1 и R^2 жирных кислот — это длинные углеводородные цепочки, которые могут быть насыщенными либо ненасыщенными. Некоторые их этих групп представлены в табл. 4.2. Группа X вместе с фосфатной частью образует полярную головку, которая в зависимости от заряда X может быть в целом заряжена или не заряжена. Некоторые из этих групп перечислены в табл. 4.3.

Свои названия фосфолипиды получают в зависимости от замещающих групп R^i и X. Например, если $R^1COOH = R^2COOH$ — это пальмитиновая кислота, а X — холин, то фосфолипид называется дипальмитоилфосфатидилхолином. Фосфолипиды, содержащие холин, иногда называют лецитинами.

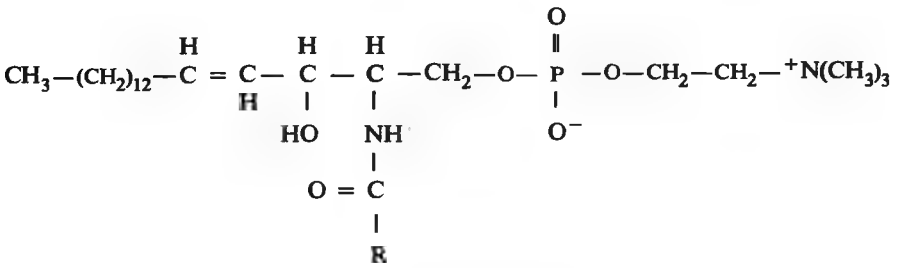
У некоторых фосфолипидов длинные алифатические цепи присоединены к глицерофосфату не через эфирную связь, хотя общая амфифильная структура у них остается прежней. Ниже приведены некоторые примеры такого рода.

Таблица 4.2

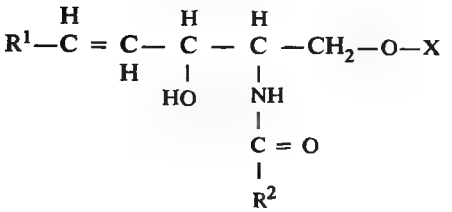
НЕКОТОРЫЕ ПРИРОДНЫЕ ЖИРНЫЕ КИСЛОТЫ[1)]

Число атомов углерода	Структура	Название кислоты: систематическое	Название кислоты: тривиальное
	Насыщенные жирные кислоты		
12	$CH_3(CH_2)_{10}COOH$	Додекановая	Лауриновая
14	$CH_3(CH_2)_{12}COOH$	Тетрадекановая	Миристиновая
16	$CH_3(CH_2)_{14}COOH$	Гексадекановая	Пальмитиновая
18	$CH_3(CH_2)_{16}COOH$	Октадекановая	Стеариновая
20	$CH_3(CH_2)_{18}COOH$	Эйкозановая	Арахиновая
24	$CH_3(CH_2)_{22}COOH$	Тетракозановая	Лигноцериновая
	Ненасыщенные жирные кислоты (если иное не указано, то двойные связи находятся в *цис*-конфигурации)		
16	$CH_3(CH_2)_5CH=CH(CH_2)_7COOH$	9-Гексадеценовая	Пальмитоолеиновая
18	$CH_3(CH_2)_7CH=CH(CH_2)_7COOH$	*цис*-9-Октадеценовая	Олеиновая
18	$CH_3(CH_2)_7CH\underset{(транс)}{=\!=}CH(CH_2)_7COOH$	*транс*-9-Октадеценовая	Элаидиновая
18	$CH_3(CH_2)_4CH=CHCH_2CH=CH(CH_2)_7COOH$	*цис, цис*-9,12-Октадекадиеновая	Линолевая
18	$CH_3CH_2CH=CHCH_2CH=CHCH_2CH=CH(CH_2)_7COOH$	9,12,15-Октадекатриеновая	α-Линоленовая
20	$CH_3(CH_2)_4CH=CHCH_2CH=CHCH_2CH=CHCH_2CH=CH(CH_2)_3COOH$	5,8,11,14-Эйкозатетраеновая	Арахидоновая
19	$CH_3(CH_2)_5HC\!-\!CH(CH_2)_7COOH$ (C, H_2)		Лактобацилловая
19	$CH_3(CH_2)_7CH(CH_3)(CH_2)_8COOH$		Туберкулостеариновая
24	$CH_3(CH_2)_{21}CH(OH)COOH$		Цереброновая

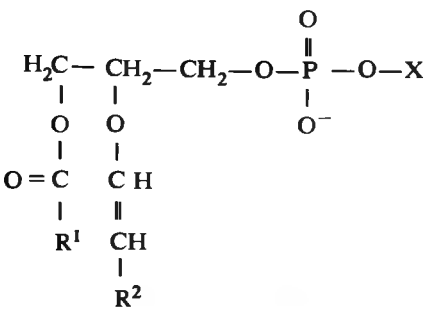
[1)] A. L. Lehninger, Biochemistry (New York: Worth, 1975), p. 281; A. White, P. Handler, E. L. Smith, R. L. Hill, I. R. Lehman, Principles of Biochemistry (New York: McGraw-Hill, 1978), pp. 39 — 40.



Сфингомиелин

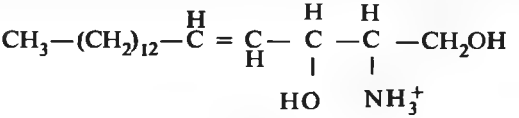


Гликолипид (X — сахарид)



Плазмалоген

В случае гликолипидов группа X — это углеводный мономер или олигомер. Сфингомиелин и гликолипиды являются производными не глицерина, а сфингозина:

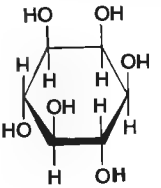
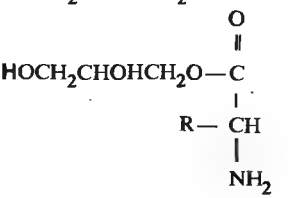
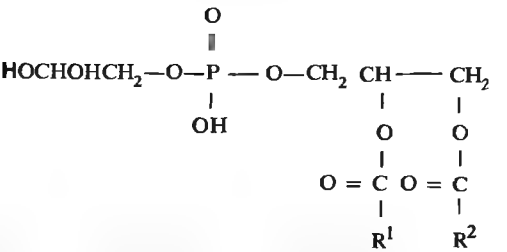


Сфингозин

Жирнокислотная цепь присоединяется с помощью амидной, а не эфирной связи. С другой стороны, плазмалогены являются производными глицерина, но одна из связей между алифатической цепью и глицериновой частью молекулы представляет собой α,β-ненасыщенную эфирную связь.

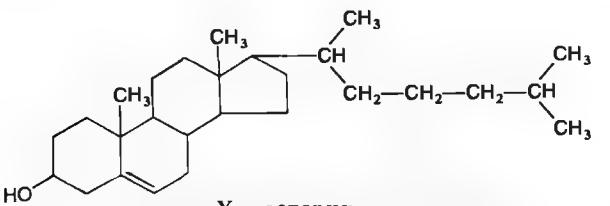
Таблица 4.3.

СПИРТЫ, ВХОДЯЩИЕ В КАЧЕСТВЕ ПОЛЯРНЫХ Х-ГРУПП В СОСТАВ НАИБОЛЕЕ РАСПРОСТРАНЕННЫХ ФОСФОГЛИЦЕРИДОВ[1)]

Фосфоглицерид	Спиртовой компонент
Фосфатидилэтаноламин	$HOCH_2CH_2NH_2$
Фосфатидилхолин	$HOCH_2CH_2N(CH_3)_3$
Фосфатидилсерин	$HOCH_2CHNH_2COOH$
Фосфатидилинозит	
Фосфатидилглицерин	$HOCH_2CHOHCH_2OH$
Фосфатидил-3′-О-аминоацилглицерин	
Кардиолипин	

[1)] Атом водорода, вместо которого присоединяется фосфат, выделен полужирным шрифтом.

Другой основной компонент мембран — это холестерин и его производные.



Холестерин

Гидроксильная группа иногда соединяется с жирной кислотой эфирной связью. Холестерин и его эфиры являются очень слабыми амфифильными соединениями.

Таблица 4.4

ЛИПИДНЫЙ СОСТАВ НЕКОТОРЫХ БИОЛОГИЧЕСКИХ МЕМБРАН[1,2)]

Липид	Эритроциты человека	Миелин человека	Митохондрии из сердечной мышцы быка	*E.coli*
Фосфатидная кислота	1,5	0,5	0	0
Фосфатидилхолин	19	10	39	0
Фосфатидилэтаноламин	18	20	27	65
Фосфатидилглицерин	0	0	0	18
Фосфатидилинозит	1	1	7	0
Фосфатидилсерин	8,5	8,5	0,5	0
Кардиолипин	0	0	22,5	12
Сфингомиелин	17,5	8,5	0	0
Гликолипиды	10	26	0	0
Холестерин	25	26	3	0

[1)] C. Tanford, The Hydrophobic Effect (New York: Wiley, 1973), p. 97.
[2)] Величины в таблице — это процентное содержание (по весу) различных липидов от общего их количества.

В табл. 4.4 приведен липидный состав некоторых мембран различного происхождения. Хотя количества отдельных компонентов и изменяются в широких пределах, преобладают двухцепочечные амфифильные фосфолипиды. Таким образом, не удивительно, что синтетические мембраны, полученные из одного чистого двухцепочечного амфифильного фосфолипида, обладают характерными свойствами природных биологических мембран.

ЧИСТЫЕ ЛИПИДЫ В БИСЛОЯХ

Важнейшим физическим свойством двухцепочечных амфифильных молекул является их склонность образовывать двойные слои в водных дисперсиях, имеющие вид отдельных пузырьков или ламеллярных (пластинчатых) структур. Такие структуры изображены на рис. 4.12; амфифильные молекулы в них объединяются так, что полярные головки контактируют с водной фазой, а углеводородные хвосты собираются вместе, образуя безводную фазу. Бислой играет роль мембраны, отделяя содержимое пузырька от внешней среды. Ясно, что толщина бислоя зависит от длины и жесткости углеводородных липидных хвостов.

В противоположность этому одноцепочечные амфифильные молекулы стремятся образовывать в водной среде глобулярные мицеллы, а не бислойные структуры. Типичная мицелла содержит до 100 молекул; ее строение показано на рис. 4.12. Таким образом, тенденция к формированию мембраноподобных бимолекулярных слоев — это специфическое свойство двухцепочечных амфифильных молекул. Для того чтобы понять причины, лежащие в основе образования мицелл или бислоев, и лучше представить их особенности, необходимо рассмотреть ряд конкретных термодинамических и структурных аспектов вопроса (гл. 25).

Плоский бислой, образованный одним-единственным типом липидов, имеет две границы (поверхности) раздела бислой — растворитель. Если раствор с обеих сторон бислоя

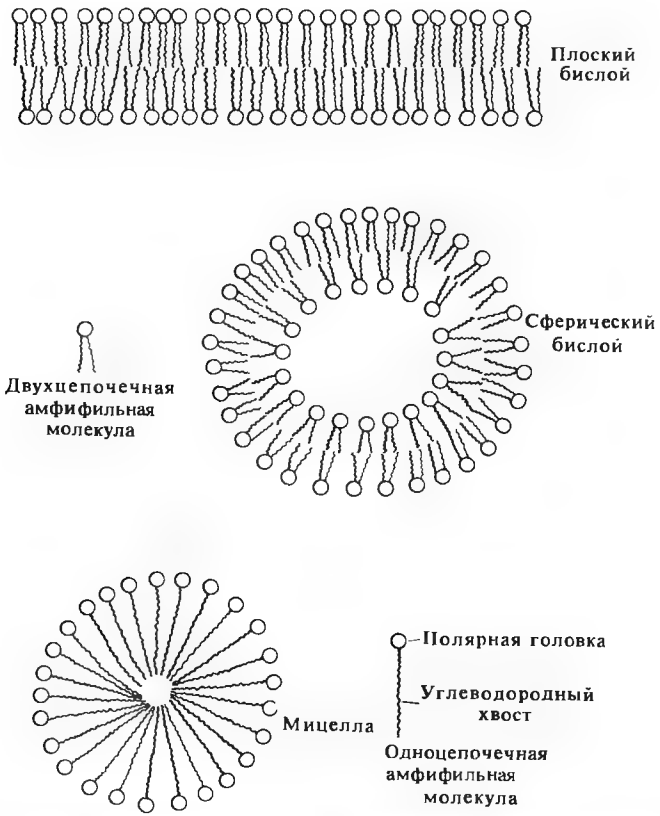


РИС. 4.12. Схематическое изображение некоторых амфифильных структур. Плоский бислой и сферический бислой (пузырек) формируются из двухцепочечных амфифильных молекул. Мицеллу образуют одноцепочечные амфифильные молекулы.

одинаков, можно ожидать, что свойства и поведение обеих поверхностей тоже не будут различаться. В противном случае может возникнуть асимметрия.

Сферический бислой достаточно большого диаметра должен вести себя практически так же, как и плоский бислой. У бислойных же пузырьков малого диаметра может возникнуть асимметрия, даже если их содержимое не отличается от внешнего раствора. Это происходит потому, что, поскольку кривизна внутренней и наружной поверхностей бислоя неодинакова, упаковка липидов на двух поверхностях несколько различается, и это может отразиться на свойствах поверхностей.

СМЕСИ ЛИПИДОВ В БИСЛОЯХ

Когда бислой состоит из липидов нескольких типов, может образоваться более сложная структура. На рис. 4.13 схематически изображено несколько сферических бислоев, состоящих из липидов двух типов. Наиболее простая структура, которую можно себе представить, — это однородный сферический бислой (рис. 4.13,*А*); обе поверхности его имеют один и тот же липидный состав, и липиды разного типа распределены между ними случайным образом.

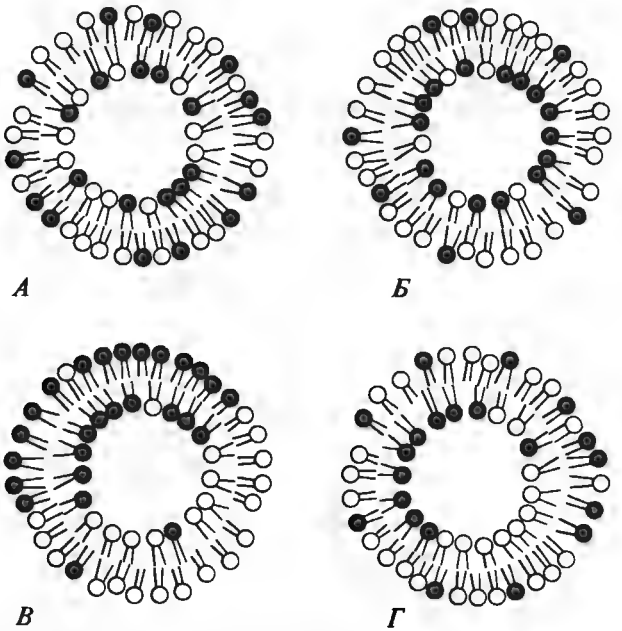


РИС. 4.13. Схематическое изображение возможных сферических бислоев, состоящих из липидов двух разных типов. *А*. Однородный бислой. *Б*. Неравномерное распределение липидов между слоями. *В*. Латеральное фазовое разделение на обеих поверхностях. *Г*. Латеральное фазовое разделение только на внутренней поверхности.

Однако из-за искривления бислоя или из-за того, что растворы внутри и снаружи пузырьков различаются, липиды могут распределиться между двумя поверхностями неравномерно (рис. 4.13,*Б*). Такая асимметрия может быть результатом разной термодинамической устойчивости данного типа липидов в монослоях с разной кривизной при данном составе раствора. Иногда она возникает и кинетически в живой клетке из-за анизотропии в переносе липидов через бислой: иапример, липиды могут специфически встраиваться во внутреннюю поверхность и удаляться с наружной. В гл. 25 будет показано, что скорость спонтанного перемещения липидов с внутренней поверхности бислоя на наружную очень мала. Таким образом, у клетки есть возможность поддержинать метастабильное распределение липидов с помощью их непрерывного встраивания и удаления.

В липидных бислоях асимметрия может возникать и еще по одной причине — вследствие латерального фазового разделения (рис. 4.13,*В,Г*). Это двумерный аналог фазовых переходов в растворе, таких, как кристаллизация или выпадение осадка. Фазовое разделение в принципе может происходить как на одной, так и на обеих поверхностях пузырька. Фазы не обязательно должны быть чистыми липидными компонентами, так же как и фазы металлических сплавов могут состоять из одинаковых атомов металлов или из их смесей.

Чтобы различить случаи равномерного и неравномерного распределения липидов между двумя поверхностями бислоя, необходимо располагать методами, которые позволили бы определять местоположение липидов. Важно уметь распознавать участки мембраны с гомогенным или гетерогенным липидным составом. Такие методы сейчас только разрабатываются (гл. 25).

ФАЗОВЫЕ ПЕРЕХОДЫ В ЛИПИДАХ

Свойства липидных молекул в бислоях сильно зависят от температуры. При достаточно низких температурах липидный слой подобен твердому телу. Липидные молекулы почти не перемещаются в плоскости мембраны, и их свойства во многих отношениях напоминают свойства углеводородных кристаллов. По мере повышения температуры часто наблюдается резкое изменение свойств бислоев, причем это может происходить при одной или нескольких температурах. Эти изменения происходят в таком узком температурном интервале, что напоминают фазовые переходы, которые наблюдаются при плавлении твердого тела. Выше температуры перехода бислой по своему поведению больше похож на жидкость. Липидные молекулы в этом случае способны быстро перемещаться в плоскости мембраны (латеральная диффузия). Температуры, при которых происходят переходы, и число этих переходов зависят от липидного состава. При высокой концентрации ненасыщенных жирных кислот образуются более жидкие бислои, обладающие более низкими температурами переходов. Иногда фазовые переходы сопровождаются латеральным фазовым разделением, описанным выше. Известны случаи, когда они приводят к затвердеванию всей мембраны.

4.4. Белки в биологических мембранах

БЕЛКОВЫЕ, ЛИПИДНЫЕ И УГЛЕВОДНЫЕ КОМПОНЕНТЫ

Имеющиеся данные позволяют полагать, что фосфолипидный бислой является структурной основой биологических мембран. Однако на самом деле биологические мембраны существенно сложнее, чем просто бислой, поскольку в них присутствуют не только липиды; в частности, значительную их часть составляют белки.

В табл. 4.5 представлен состав некоторых клеточных мембран. Все мембраны содержат белки, липиды и углеводы. Углеводы во всех случаях составляют по весу небольшую часть мембран — как правило, 10% или меньше. Они связаны с основными компонентами мембран — липидами и белками, т.е. обычно присутствуют в виде гликолипидов и гликопротеидов.

Содержание как белков, так и липидов в мембранах сильно варьирует в зависимости от типа клеток. Например, в миелиновых мембранах отношение количества белков к количеству липидов составляет около 0,2, в то время как во внутренней мембране митохондрий это отношение в 15 раз выше. Такая вариабельность несомненно отражает разнообразие функций биологических мембран. Так, миелин, по-видимому, в основном играет роль изолятора и не требует присутствия больших количеств белков. С другой стороны, внутренняя мембрана митохондрий участвует в ферментативных и транспортных процессах, которые протекают с помощью множества разных белков.

ТИПЫ И РАСПОЛОЖЕНИЕ БЕЛКОВ

С. Сингер выдвинул положение о том, что мембранные белки можно подразделить на две большие группы, периферические и интегральные, в зависимости от того, как они связаны с мембраной (рис. 4.14). *Периферические* белки экстрагируются из мембраны при сравнительно мягкой обработке, например при повышении ионной силы. Эти белки обычно устойчивы в водных растворах и не очень прочно связаны с липидами. Примерами белков этого типа могут служить цитохром *c* из митохондриальной мембраны и спектрин (белковый компонент мембраны эритроцита).

Таблица 4.5

СОСТАВ НЕКОТОРЫХ КЛЕТОЧНЫХ МЕМБРАН[1)]

Мембрана	Белки, %	Липиды, %	Углеводы, %	Отношение количества белков к количеству липидов (по весу)
Миелин	18	79	3	0,23
Плазматические мембраны				
кровяные пластинки	33 — 42	58 — 51	7,5	0,7
клетки печени мыши	46	54	2 — 4	0.85
эритроцит человека	49	43	8	1,1
амеба	54	42	4	1,3
клетки печени крысы	58	42	(5 — 10)[2)]	1,4
L-клетки	60	40	(5 — 10)[2)]	1,5
клетки HeLa	60	40	2,4	1,5
ядерные мембраны клеток печени крысы	59	35	2,9	1,6
палочки сетчатки из глаза быка	51	49	4	1,0
наружная митохондриальная мембрана	52	48	(2 — 4)[2)]	1,1
Саркоплазматический ретикулум	67	33	—	2,0
Ламеллы хлоропластов шпината	70	30	(6)[2)]	2,3
Внутренняя митохондриальная мембрана	76	24	(1 — 2)[2)]	3,2
Грамположительная бактерия	75	25	(10)	3,0
Пурпурная мембрана *Halobacterium*	75	25	—	3,0
Микоплазма	58	37	1,5	1,6

[1)] G. Guidotti, Ann. Rev. Biochem., **41**, 731 (1972).
[2)] Получено косвенным путем.

Интегральные белки, напротив, с трудом поддаются экстракции. К примеру, для их удаления необходимо обработать мембрану органическими растворителями. Эти белки часто выделяются в комплексе с липидами, без которых в водной среде они обычно либо слипаются, либо выпадают в осадок. Наконец, интегральные белки обычно сильно гетерогенны по размеру, и во многих мембранах они составляют более 70% общего количества белка.

Эти существенные различия в свойствах и поведении интегральных и периферических мембранных белков находят свое отражение в таком важном параметре, как их аминокислотный состав. Как показано в гл. 2, зная аминокислотный состав белка, можно с уверенностью сказать, является ли он интегральным мембранным белком.

Естественно полагать, что большинство периферических белков расположено на поверхности мембран. Эти белки могут находиться как на одной из поверхностей, так и с обеих сторон мембраны. Вполне правдоподобным кажется предположение, что во многих случаях они связаны с мембраной через интегральные белки, а не путем прямого контакта с липидами.

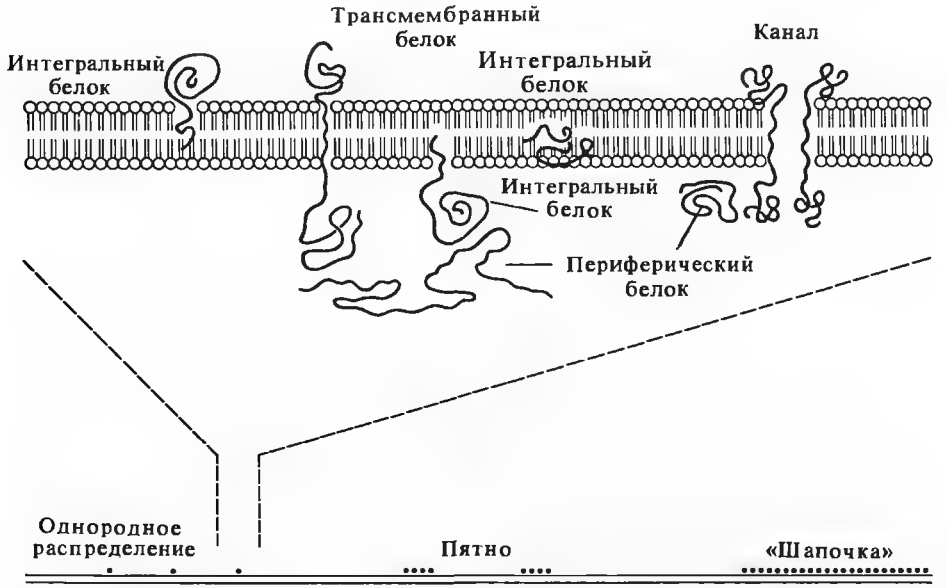


РИС. 4.14. Типы и расположение мембранных белков. Вверху при большом увеличении изображена часть структуры, представленной в нижней половине рисунка.

ПОВЕРХНОСТНАЯ ЛОКАЛИЗАЦИЯ БЕЛКОВ

Интегральные белки могут быть локализованы на какой-то одной или на обеих поверхностях мембраны, а иногда они даже пронизывают мембрану. Эти последние белки называются трансмембранными. По-видимому, интегральные белки имеют участок, насыщенный гидрофобными остатками, что позволяет им вступать в прямой контакт с липидами. У некоторых интегральных белков, расположенных на какой-то одной поверхности, есть гидрофобный хвост, посредством которого они прикрепляются к мембране. Если с помощью ферментов отсечь этот хвост, то он останется в мембране, в то время как головка перейдет в раствор и будет вести себя практически так же, как и обычный растворимый белок. Есть основания считать, что у трансмембранных белков гидрофобная последовательность находится где-то в середине молекулы. Примером белка такого рода является гликофорин — трансмембранный белок эритроцитов. Его аминокислотная последовательность приведена на рис. 2.16.

Иногда установить поверхностную локализацию мембранных белков позволяют довольно простые эксперименты, основанные на использовании реагентов, которые не могут проходить через липидный бислой (например, антител, вырабатываемых против определенного мембранного белка). Взяв замкнутый сферический пузырек или целую клетку, определяют, связывается ли с данным белком реагент, добавленный во внешний раствор. Затем пузырек разрушают и смотрят, связывается ли теперь реагент с белками, которые исходно были локализованы на внутренней поверхности мембраны. Кроме того, существуют методы образования пузырьков в присутствии заблокированного реагента. После того как пузырьки образовались, реагент из внешней среды отмывают, затем заблокированный реагент активируют, и он может атаковать белки, находящиеся на внутренней поверхности.

С помощью метода поверхностной метки довольно трудно отличить белки, локализованные на обеих поверхностях, от трансмембранных белков. Для этого необходимо определить, доступна ли одна и та же полипептидная цепь с обеих сторон мембраны одновременно или нет.

ЛАТЕРАЛЬНОЕ РАСПОЛОЖЕНИЕ БЕЛКОВ

Для детального описания мембранных белков требуется знать не только то, на какой поверхности они находятся, но и их распределение по этой поверхности. Белки могут перемещаться в мембране независимо, а могут быть связаны с другими белками. Они способны образовывать симметричные олигомеры, спиральные или пластинчатые структуры. О четвертичной структуре мембранных белков почти ничего не известно. Этот вопрос вызывает особый интерес в связи с возможностью образования каналов в мембране либо при участии отдельных молекул белка, имеющих внутреннюю полость, либо в результате ассоциации белков с образованием структуры с полостью посередине. Предположение о наличии каналов возникает всякий раз, когда мембрана становится хорошо проницаемой для ионов, так как чистые липидные бислои являются хорошими изоляторами. Однако о механизме образования каналов и о регуляции их проницаемости на молекулярном уровне известно очень мало.

Ближайшие соседи данного мембранного белка могут быть определены на молекулярном уровне с помощью методов, аналогичных тем, что применяются для установления четвертичной структуры обычных белков (гл. 2, 8 и 14). При макроскоцическом же описании необходимо знать, распределен ли каждый тип белков в мембране равномерно или проявляет какие-либо признаки латерального фазового разделения. Для исследования дальнего порядка в организации мембранных белков особенно полезна электронная микроскопия. В одних системах белки могут быть распределены равиомерно, в других — образовывать пятна и даже «шапочки», которые покрывают значительную часть поверхности клетки (рис. 4.14).

ЭЛЕКТРОННАЯ МИКРОСКОПИЯ МЕМБРАН, ОБРАБОТАННЫХ МЕТОДОМ ЗАМОРАЖИВАНИЯ—ТРАВЛЕНИЯ

За последние десять лет благодаря использованию электронного микроскопа и разработке методов замораживания—скалывания и замораживания—травления в изучении строения мембран достигнуты значительные успехи. В основе указанных методов лежит быстрое замораживание образца до очень низких температур, при котором в нем, как полагают, происходят лишь минимальные изменения. Замороженный образец раскалывают ножом, с тем чтобы обнажить его внутреннюю поверхность, и готовят для электронной микроскопии.

На рис. 4.15 представлены основные этапы этой процедуры в соответствии с методикой Д. Брантона и др. Как показано на рисунке, образец сначала быстро замораживают в жидком фреоне, а затем производят скол охлажденным ножом в вакууме (чтобы избежать обмерзания ножа). Травление предстанляет собой испарение некоторой части льда в вакууме при температуре около -100 °С. Обнажившуюся после травления поверхность оттеняют платиной и углеродом и далее снимают образовавшуюся пленку — реплику поверхности, которую можно исследовать с помощью электронного микроскопа.

Установлено, что, как правило, скол в мембране проходит вдоль поверхности, разделяющей два углеводородных слоя (рис. 4.16). Таким образом, именно эта внутренняя область подвергается травлению и именно с нее снимают реплику. На рис. 4.17 представлена фотография скола мембраны из *Acholeplasma laidlawii*; видна относительно гладкая по-

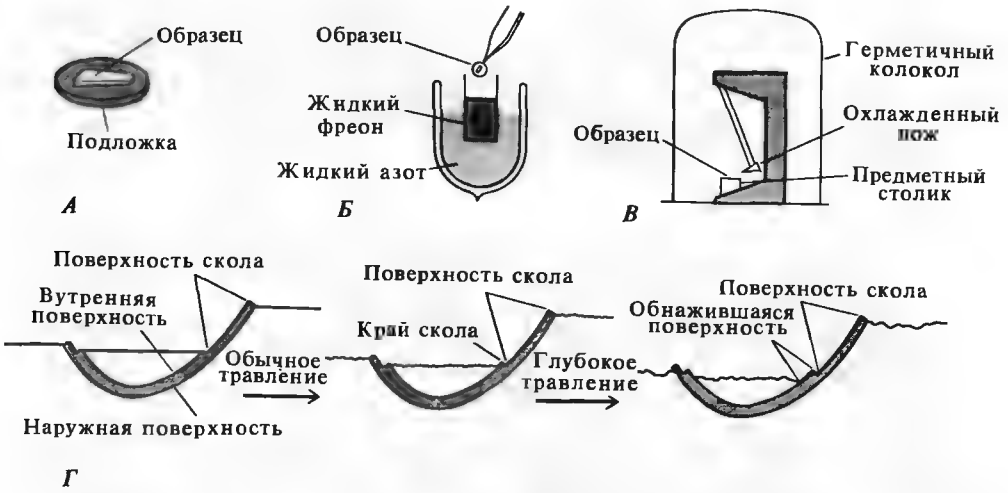


РИС. 4.15. Методнка замораживания—травления. Образец помещают на подложку (*А*), замораживают (*Б*) н в вакууме делают скол охлаждениым ножом (*В*). При травлении (*Г*) лед сублимируется, и обнажаются поверхностные структуры. [D. Branton, Proc.Nat. Acad.Sci. USA, **55**, 1048, (1966).]

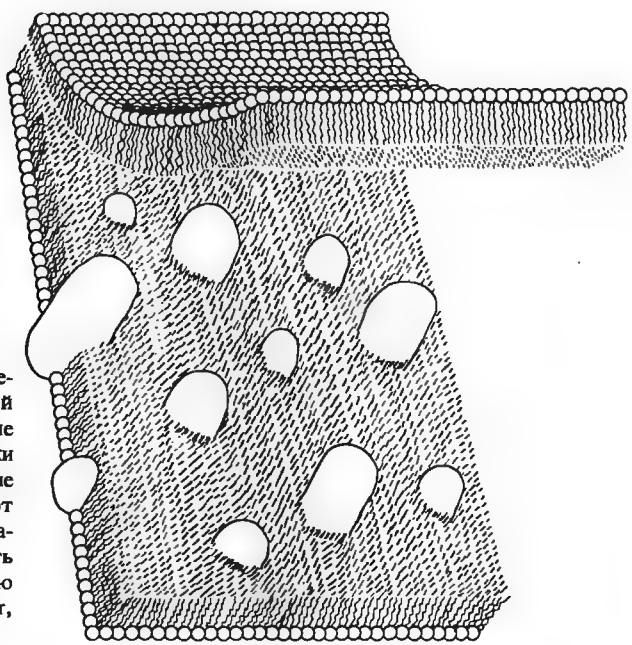


РИС. 4.16. Схематическое изображение поверхности скола, наблюдаемой под электронным микроскопом после замораживания—скалывания. Белки могут остаться на любой половине липидного бислоя, и в зависимости от того, с какой из этих половин снимается отпечаток, онн будут выглядеть как выступающие над поверхностью частицы илн как вмятины. [S.J.Singer, Hosp. Pract., **8**, 81 (1973).]

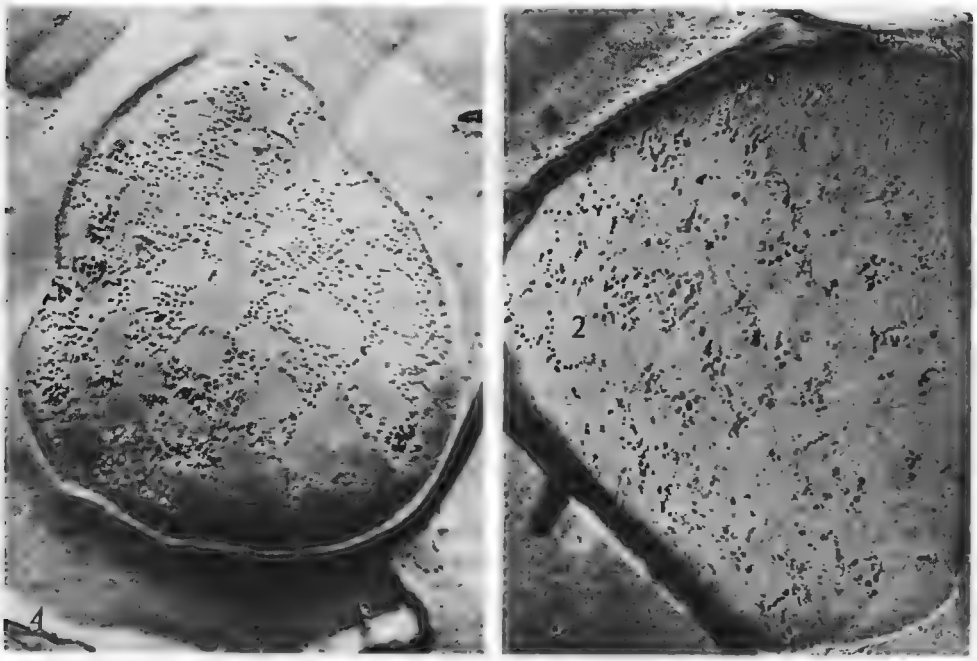


РИС. 4.17. Фотография скола мембраны из *Acholeplasma laidlawii*. 1 — выпуклая поверхность скола, *2* — вогнутая поверхность скола. [A.J. Verkleij et al., Biochim. Biophys. Acta, 288, 326 (1972).]

верхность скола, испещренная множеством практически случайно распределенных частиц. Скорее всего этими частицами являются белки или их агрегаты. Картина, представленная на рис. 4.17, довольно типична для мембран, находившихся до замораживания при физиологической температуре, когда липиды обычно находятся в жидком состоянии (выше температуры перехода в липидной мембране). У мембран, хранившихся до замораживания при температуре ниже температуры перехода, на поверхности скола наблюдаются гребни, а частицы образуют большие кластеры.

Гребни можно наблюдать и у чисто липидных бислойных пузырьков (липосом), которые были заморожены в состоянии ниже температуры фазового перехода; в то же время поверхность скола всегда была гладкой, если вначале липосомы находились при температуре выше температуры перехода. Таким образом, липосомы при замораживании—скалывании ведут себя во многом так же, как природные мембраны. Это, как и многое другое, означает, что структурной основой биологических мембран является такой же липидный бислой, что и у липосом. Фотография на рис. 4.17 (и другие ей подобные) показывает, что в природных мембранах белки рассеяны по бислойному матриксу, образованному фосфолипидами.

РАСПРЕДЕЛЕНИЕ ОТДЕЛЬНЫХ МАКРОМОЛЕКУЛ, ПОЛУЧЕННОЕ С ПОМОЩЬЮ ЭЛЕКТРОННОГО МИКРОСКОПА

Электронная микроскопия может быть использована также и для выявления отдельных мембранных *поверхностных* компонентов. Для этого необходимо вначале «окрасить» интересующие нас молекулы, так чтобы они стали видны под электронным микро-

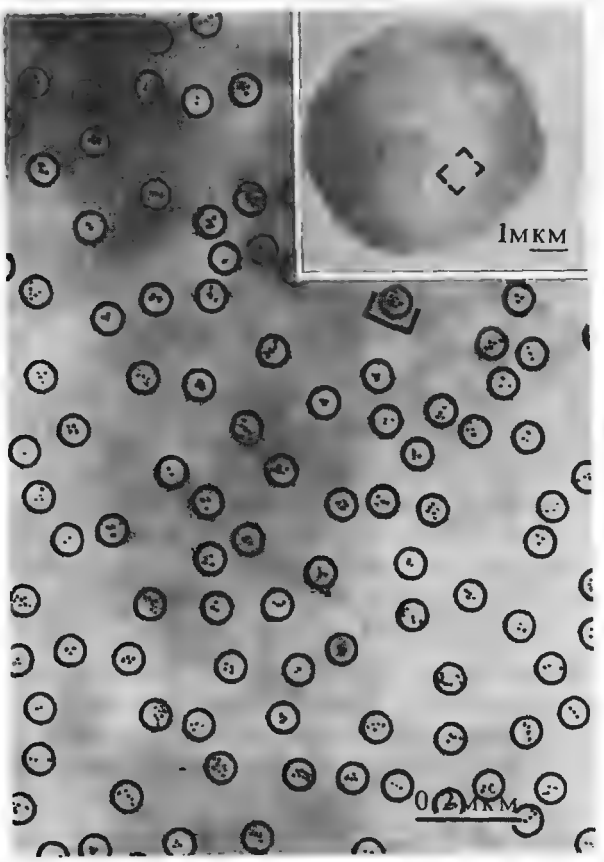


РИС. 4.18. Распределенне $Rh_0(D)$-антигенов на поверхности эритроцита человека. Каждый заключенный в кружок кластер черных точек соответствует одной молекуле Rh_0D. Черные точки — молекулы ферритина, связанные с козьими антителами; козьи антитела прореагировали с антителамн человека, которые в свою очередь связаны с молекулами Rh_0 (D). [G.L. Nicolson et al., Proc. Nat. Acad. Sci. USA, **68**, 1416 (1971).]

скопом. С этой целью используют электроноплотные метки, связанные с реагентом, который специфически взаимодействует с исследуемым компонентом.

Пример такого подхода дает работа Сингера и др., в которой исследонано распределение антигена $Rh_0(D)$ на поверхности мембраны эритроцита человека. Сначала O,Rh-положительные клетки были обработаны антителами человека, выработаниыми против специфического антигена. После того как антитела связались со специфическими поверхностными антигенами, клетки были разрушены, а их мембраны распластаны на сетке электронного микроскопа. Затем распластанные мембраны были обработаны *козьими* антителами IgG, выработанными против *человеческих* антител IgG, связанных с антигенами $Rh_0(D)$ на мембране. (Таким образом, связанные антитела человека выступали в качестве антигенов для козьих антител.) Козьи антитела были предварительно связаны с ферритином — белком, богатым железом, который обладает достаточно большой электронной плотностью, чтобы его легко можно было различить под электронным микро-

скопом. Таким образом, ферритин, связанный с козьими антителами, является тем особым красителем, который делает поверхностные антигены $Rh_0(D)$ видимыми.

Полученные результаты представлены на рис. 4.18; видно распределение ферритина (черные точки) на наружной поверхности мембраны. От двух до восьми точек образуют кластеры, заключенные на рисунке в кружки. Каждый кластер соответствует одиому антигену $Rh_0(D)$, а каждая точка — одному козьему антителу, прореагировавшему с человеческим IgG, который связан с этим антигеном. Рис. 4.18 показывает, что поверхностный антиген $Rh_0(D)$ распределен по наружной поверхности мембраны случайным образом. Это согласуется с данными, полученными методом замораживания—травления: как видно из рис. 4.17, белковые частицы распределены по внутренней поверхности скола мембраны *Acholeplasma laidlawii* нерегулярно.

ОБЩАЯ КАРТИНА СТРОЕНИЯ БИОЛОГИЧЕСКИХ МЕМБРАН

Полезно рассмотреть рабочую модель биологической мембраны. Схематическая иллюстрация на рис. 4.19 отражает основные черты жидкостно-мозаичной модели, предложенной С. Сингером. Фосфолипидный бислой можно представить как жидкий матрикс толщиной от 60 до 100 Å, в который погружены различные интегральные белки. Одни белки пронизывают всю толщу мембраны, другие локализованы на одной из ее сторон. В результате мембрана является асимметричной — факт, которому есть масса подтверждений. Белковые молекулы имеют как гидрофобные, так и гидрофильные участки, что позволяет им термодинамически наиболее выгодно встраиваться в соответствующие места фосфолипидного бислоя. Дальний порядок в распределении каждого данного белка, повидимому, отсутствует, о чем говорит распределение антигена $Rh_0(D)$ (рис.4.18). С другой стороны, для каждого данного белка может наблюдаться ближний порядок.

Мембрана — это динамическая система, поскольку белки могут совершать в ней латеральные перемещения. На такие перемещения указывает множество экспериментальных данных. Возможно также, что белковые молекулы вращаются вокруг осей — перпенди-

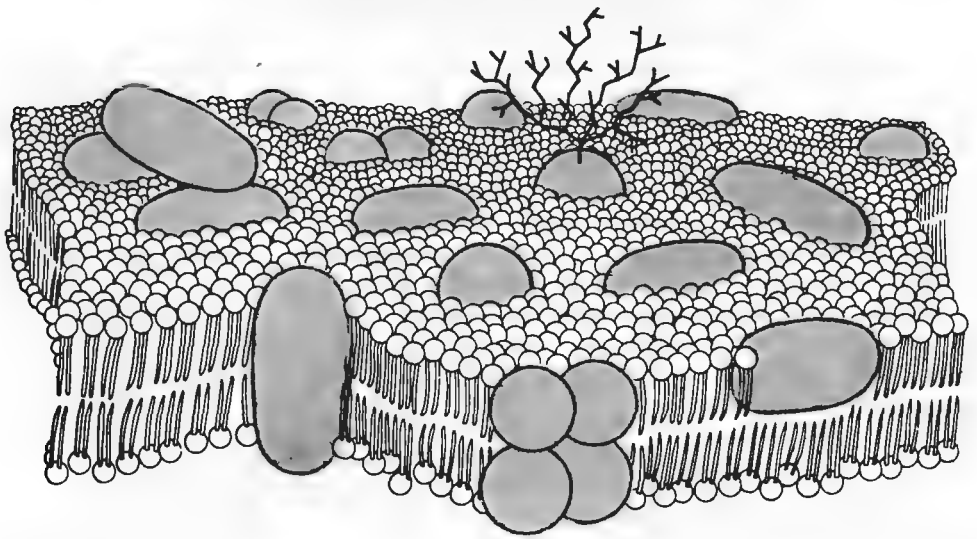


РИС. 4.19. Обобщенная модель биологических мембран. [J.M. Clark, Jr., R.L. Switzer, Experimental Biochemistry, 2nd ed. (San Francisko: W.H. Freeman and Company. 1977).]

кулярных и параллельных плоскости бислоя. Такое вращение может играть роль при выполнении многих функций, присущих мембране. Однако конкретные исследования этих аспектов только начинаются..

МЕМБРАНА ЭРИТРОЦИТА

Плазматическая мембрана эритроцита — это наиболее сложная природная мембрана, изученная сравнительно детально. Основные полипептиды, обнаруженные в ней, представлены на рис. 4.20, *А*.

Около 40% массы мембраны составляют периферические белки. Это компоненты 1; 2; 4,1; 4,2; 5 и 6 (номера соответствуют полосам, наблюдаемым при электрофорезе). Все они находятся на внутренней поверхности мембраны, на что указывают энзимологические и иммунохимические тесты. Компоненты 1 и 2 — это высокомолекуляриый фибриллярный белок, называемый спектрином. Отдельные его полипептидные цепи имеют мол. массу в пределах от 225000 до 250000. Эти компоненты, по-видимому, тесно взаимодействуют друг с другом, поскольку между ними легко образуются сшивки.

Белок полосы 5 — это актин или актиноподобный белок. Он имеет мол. массу 45000 и может образовывать гомополимеры. У многих исследователей сложилось впечатление, что спектриновые и актиновые компоненты связаны между собой и образуют фибриллярную сетку, выстилающую всю клеточную мембрану (что схематически показано на

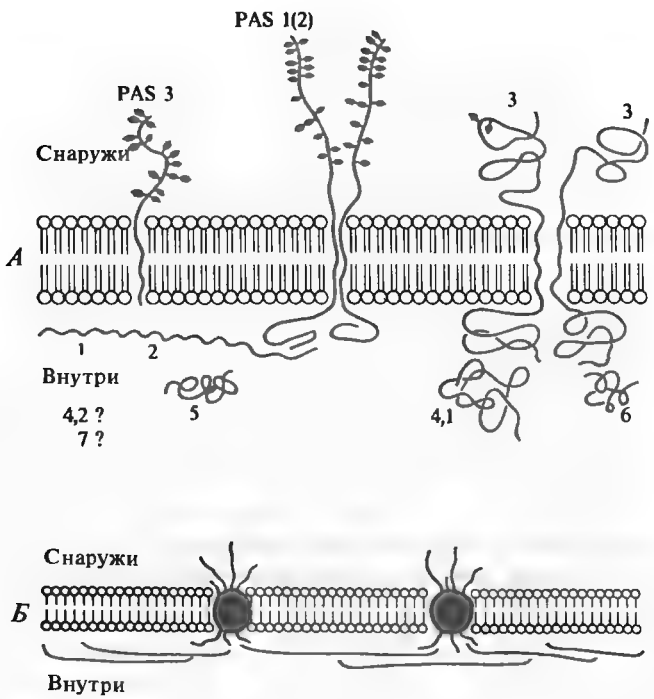


РИС. 4.20. Возможная модель мембраны эритроцита. *А*. Основные полипептиды (см. текст). *Б*. Возможная структура внутренней фибриллярной белковой сети. [V.T. Marchesi et al., Ann. Rev. Biochem., **45**, 667 (1976).]

рис. 4.20,*Б*). Возможно, эта сетка играет роль в поддержании формы эритроцита, а также в определении местоположения тех или иных мембранных белков.

Белок полосы 6 — глицеральдегид-3-фосфат—дегидрогеназа. Этот белок изображен на рис. 4.20,*А* рядом с белком полосы 3 (интегральным), но некоторые недавно полученные данные позволяют считать, что на самом деле он связан только со спектрином. О полосах 4,1 и 4,2 почти ничего не известно.

Белок полосы 3 и PAS-1 — это наиболее важные интегральные мембранные белки. Компонент 3 — гликопротеид, содержащий примерно от 5 до 8% углеводов. Он изображен в виде димера, так как в относительно мягких условиях в нем образуются поперечные сшивки. На основании данных по химической модификации можно полагать, что компонент 3 — трансмембранный белок. Он может играть роль в переносе анионов или глюкозы.

PAS-1 — это гликофорин, свойства которого широко обсуждались в данной главе и в гл. 2. Он изображен в виде димера, так как оказалось, что он ведет себя как димер в растворах детергентов, например додецилсульфата натрия. Не исключено, что гликофорин составляет основную массу глобулярных структур, которые наблюдаются на сколе мембраны эритроцита при исследовании методом замораживания—скалывания. Существуют некоторые указания на то, что участки молекулы гликофорина, проникающие внутрь клетки, непосредственно взаимодействуют со спектриновой сеткой (как показано на рис. 4.20,*Б*).

PAS-3 — последний существенный мембранный компонент. Это гликопротеид с пептидной составляющей, имеющей мол. массу примерно 25000. Вероятно, PAS-3 — полуинтегральный белок, т.е. белок, легче экстрагируемый из мембраны, чем PAS-1 и белок полосы 3.

Основные липидные компоненты мембраны эритроцита перечислены в табл. 4.4. Последние данные указывают на то, что они распределены в бислое несимметрично: около 70% всего фосфатидилхолина и от 80 до 85% всего сфингомиелина находятся на наружной поверхности бислоя. Примерно 80 — 90% фосфатидилсерина и фосфатидилэтаноламина локализованы на внутренней поверхности мембраны. Имеющиеся данные говорят о том, что холестерин, последний существенный липидный компонент, распределен между двумя слоями приблизительно равномерно.

4.5. Взаимодействия между белками и липидами

ВЗАИМОДЕЙСТВИЯ МЕЖДУ БЕЛКАМИ И ЛИПИДАМИ В БИСЛОЕ

Специфические взаимодействия между белками и липидами должны определять основные особенности структуры и свойств мембраны. Исходя из общих соображений о нековалентных взаимодействиях, можно с полной уверенностью сказать, что гидрофобные хвосты липидных молекул должны охотно взаимодействовать с неполярными боковыми цепями аминокислот, а полярные головки — с полярными частями белковых молекул. Однако существенные детали этих взаимодействий еще остаются невыясненными. Перечислим несколько вопросов, исследование которых только начинается.

Вызывают ли мембранные белки какие-либо возмущения в прилегающей липидной фазе? Такие возмущения могут состоять в специфическом притяжении определенных липидных компонентов, увеличении подвижности или иммобилизации соседних липидов, а также в изменении глобальных свойств мембраны (таких, как толщина бислоя, его кривизна и проницаемость). Некоторые данные указывают на то, что поведение липидов, контактирующих с белками, отличается от поведения всех остальных липидов. Пограничные липидные молекулы, вероятно, полностью аналогичны тем молекулам воды, которые тесно связаны с растворимыми белками (см. гл. 10).

Влияют ли липиды на свойства мембранных белков? Например, будет ли мембранный белок функционировать в любом бислое или для его работы иеобходим особый липидный состав? Совершенно очевидно, что функционирование некоторых ферментов связано с текучестью мембран. Это особенно легко понять в случае реакций, для протекания которых субстраты или ферменты должны диффундировать в мембране по направлению друг к другу. Имеющиеся данные пока не позволяют ответить на вопрос, «различают» ли белки бислои с одинаковой температурой перехода, но с разным липидным составом. Отметим, однако, что некоторые белки претерпевают существенные изменения, когда температура при своем изменении проходит область фазового перехода в липидах.

Проникают ли липиды внутрь мембранного белка или связываются с его поверхностью — неясно. Здесь крайне нужны модельные системы, в которых белок взаимодействует лишь с несколькими липидами и отсутствуют те сложности, которые привносит мембрана. Весьма многообещающим представляется использование интегрального мембранного фермента *C-55*-изопреноид-алкоголь—фосфокиназы и периферического фермента пируватоксидазы (Gennis, Jonas, 1977).

В последнее время появились лишь первые данные относительно формы мембранных белков (гл. 14). Прежде чем делать достаточно общие выводы, необходимо будет определить трехмерную структуру нескольких таких белков. Весьма правдоподобным кажется предположение, что те части молекул мембранных белков, которые находятся внутри бислоя, имеют α-спиральную или β-складчатую структуру. Это наиболее эффективный способ спаривания доноров и акцепторов водородных связей.

ЛИПОПРОТЕИДЫ

Взаимодействия между белками и липидами имеют место не только в мембранах, но и в липопротеидах. Липопротеиды — это мицеллоподобные агрегаты. Здесь мы рассмотрим только липопротеиды плазмы, которые являются нековалентными агрегатами различных липидов и пептидов. Как уже отмечалось, на поверхности бактериальных клеток существуют и ковалентные липидно-белковые соединения. В плазме различают четыре основных класса липопротеидов. Это (в порядке возрастания плотности) хиломикроны, липопротеиды с очень низкой плотностью (ЛОНП), липопротеиды с низкой плотностью (ЛНП) и липопротеиды с высокой плотностью (ЛВП).

О липопротеидах первого класса известно относительно мало, но свойства липопротеидов остальных трех классов отчасти установлены; в табл. 4.6 суммированы все известные по этому поводу данные. Основная функция липопротеидов плазмы — это транспорт липидов, в частности ЛНП являются главными переносчиками холестерина. Липопротеиды, выделенные от человека и из других организмов, — это крупные глобулярные частицы, несколько различающиеся по содержанию белков и липидов. Липопротеиды всех классов содержат небольшое число главных белковых компонентов, причем аминокислотные последовательности большинства из них в ЛОНП и ЛВП известны. Для белков этих двух классов характерен необыкновенно высокий процент α-спиральности; эта структура стабилизируется благодаря присутствию липидов, так как у очищенных апопротеидов процент α-спиральности ниже, чем у природных липопротеидов.

Как и ожидалось, низкая плотность белка сочетается с высоким содержанием липидов. Высокое содержание липидов характерно также для частиц большого размера. По-видимому, в каждом классе липопротеидов преобладает какой-то один тип липидов, хотя в той или иной степени в них представлены самые разные липиды. Более половины липидов в ЛОНП — это триглицериды, в ЛНП — эфиры холестерина, в ЛВП — фосфолипиды. Заметим, что два компонента (триглицериды и эфиры холестерина) почти неполярны и нерастворимы в воде. Поэтому необходимо, чтобы остальные белки и липиды обеспечивали образование стабильной суспензии, которая могла бы переноситься током кро-

Таблица 4.6

СВОЙСТВА ЛИПОПРОТЕИДОВ ПЛАЗМЫ [1]

	Липопротеид с очень низкой плотностью	Липопротеид с низкой плотностью	Липопротеид с высокой плотностью
Плотность, г/см^3	0,95 — 1,006	1,019 — 1,063	1,063 — 1,21
Мол. масса	$5 \cdot 10^6$	$2 \cdot 10^6$	От $1{,}7 \cdot 10^5$ до $3{,}6 \cdot 10^5$
Содержание белка, % (по весу)	10	25	50
Основные белковые компоненты	Апо-С-I (57 аминокислот) Апо-С-II (78 аминокислот) Апо-С-III (79 аминокислот) Апо-Е (мол. масса 33000) Апо-В	Апо-В (мол. масса до 275 000)	Апо-А-I (245 аминокислот) Апо-$A\text{-II}_I$ (димер из двух цепей, содержащих по 77 аминокислотных остатков)
Вторичная структура белков	α-Спираль, 73 % Другие типы конформации, 27 %	α-Спираль, 25 % β-Конформация, 37 % Неупорядоченная структура, 37 %	α-Спираль, 70 % β-Конформация 11 % Неупорядоченная структура, 19 %
Вторичная структура апопротеидов	α-Спираль, 56 %		α-Спираль, 52 % β-Конформация, 10 % Неупорядоченная структура, 38 %
Содержание липидов (в процентах от суммарного веса липопротеида)			
триглицериды	55	6 — 8	Малое
фосфолипиды	20	21 — 22	>20
холестерин	10	7 — 8	Малое
эфиры холестерина	5	36 — 41	15

[1] Из результатов, обобщенных J. D. Morresett, R. L. Jackson, A. M. Gotto, Jr., Ann. Rev. Biochem., **44**, 183 (1975) и L. C. Smith, H. J. Pownall, A. M. Gotto, Jr., Ann. Rev. Biochem., **47**, 751 (1978).

ви. Большинство белков, содержащихся в ЛОНП, — это сравнительно низкомолекулярные белки. Те, что присутствуют в ЛВП, относятся скорее к белкам среднего размера, за исключением белка апо-А-II, который является весьма необычным симметричным димером, состоящим из двух одинаковых цепей, соединенных единственной дисульфидной связью. Белковые компоненты ЛНП изучены не очень хорошо, но весьма возможно, что они крупнее других аполипопротеидов.

При исследовании природных липопротеидов возникают два основных вопроса: 1) какова организация липидных и белковых компонентов в глобулярных частицах и 2) как именно взаимодействуют между собой белки и липиды? Известно здесь пока очень мало. Большинство липопротеидов, по-видимому, образуют мицеллы, внутри которых находятся неполярные липиды, а снаружи — полярные части белков и другие липиды. Однако многое при этом остается неясным. Расположены ли белковые субъединицы в мицеллах упорядоченным образом? Некоторые исследования ЛНП позволяют предположить, что это действительно так. Контактируют ли белковые субъединицы друг с другом или они

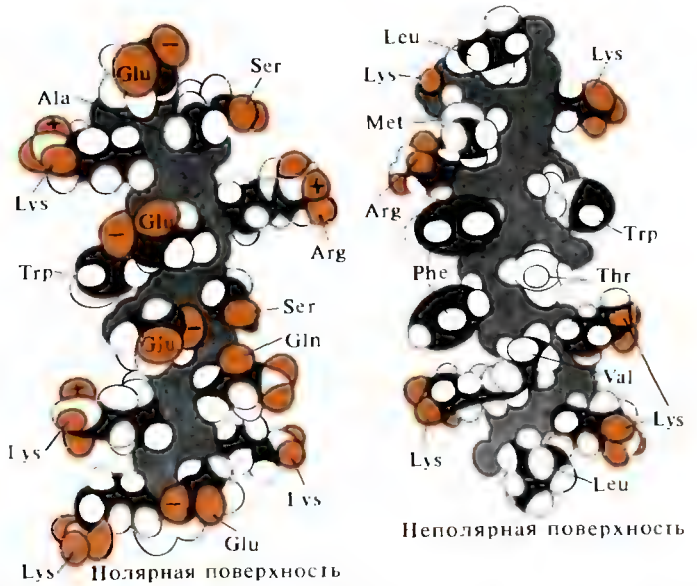


РИС. 4.21. Схематическое изображение α-спиральной структуры одного из участков пептидного компонента аполипопротеида С-I. Заряженные группы изображены в цвете. [J.D. Morrisett et al., Ann.Rev. Biochem., 44, 183 (1975).]

плавают в «липидном море» независимо? Белки апо-А-I и апо-А-II из ЛВП в отсутствие природных липидов, будучи взяты в эквимолярных количествах, образуют димер 1 : 1. Очевидно, это взаимодействие имеет место также и в присутствии липидов, так как в ЛВП эти два белка легко сшиваются 1,5-дифтор-2,4-динитробензолом.

Было показано, что по крайней мере некоторые из полипептидов в составе липопротеидов могут взаимодействовать с огромным числом липидов. Например, белок апо-С-III способен связывать от 18 до 80 молекул фосфатидилхолина. Исследования по связыванию липидов фрагментами белковых молекул, а также по защите целых молекул от протеолиза указывают на то, что основная масса липидов связывается с С-концевой частью белка. Аналогичные исследования, проведенные с белком апо-А-II, показали, что фосфолипиды связываются в основном с двумя третями молекулы белка со стороны С-конца. Тот факт, что липиды увеличивают степень α-спиральности белка, приводит к интересной гипотезе о детальной природе этого связывания. Проще всего объяснить данное явление тем, что липиды связываются именно с α-спиральными участками.

Когда аминокислотные цепочки, предположительно связывающиеся с липидами, свернуты в α-спираль, обнаруживается следующий любопытный факт: часть таких участков в белке апо-С-I (рассматриваемых как цилиндр) имеет неполярную и полярную поверхности (рис. 4.21). Поразительная особенность полярной поверхности заключается в том, что отрицательно заряженные боковые цепи лежат в ее середине, в то время как положительно заряженные группы расположены по краям. Заманчиво предположить, что фосфолипиды связаны с α-спиралями так, что отрицательно заряженный фосфат соседствует с положительно заряженным краем, а несколько метиленовых групп жирной кислоты, ближайших к фосфату, контактируют со смежной неполярной поверхностью пептида. Тогда остальная часть молекулы жирной кислоты может быть направлена наружу и взаимодействовать с другими липидными молекулами.

По-видимому, можно поставить такие эксперименты, которые позволили бы непосредственно исследовать характер взаимодействия между липидами и белками. Такие эксперименты важны не только для понимания структуры липопротеидов, но и для решения значительно более общей проблемы структуры и функций мембранных белков.

НЕВЫЯВЛЕННЫЕ МАКРОМОЛЕКУЛЯРНЫЕ КОМПЛЕКСЫ

Из четырех основных типов макромолекул — липидов, полисахаридов, белков и нуклеиновых кислот — в принципе могут образовываться шесть типов ковалентных парных комплексов и еще шесть нековалентных. Из табл. 4.7 видно, что из 12 потенциально

Таблица 4.7

РАСПРОСТРАНЕННОСТЬ ОБНАРУЖЕННЫХ К НАСТОЯЩЕМУ ВРЕМЕНИ ГЕТЕРОГЕННЫХ ПРИРОДНЫХ АССОЦИАТОВ МАКРОМОЛЕКУЛ

Ассоциат	Ковалентная связь	Нековалентная связь
Белок — нуклеиновая кислота	Редко	Очень часто
Белок — липид	» »	» »
Белок — сахарид	Очень часто	Редко
Нуклеиновая кислота — сахарид	Редко	—
Липид — сахарид	Довольно часто	—
Липид — нуклеиновая кислота	—	—

возможных комплексов среди упорядоченных долгоживущих стабильных клеточных структур пока удалось обнаружить только 8. Есть ли в природе остальные четыре — пока неизвестно. Очень настораживает то, что не удалось обнаружить ни одного упорядоченного комплекса липидов и нуклеиновых кислот. В принципе такие комплексы могли бы образоваться в шероховатом эндоплазматическом ретикулуме, который состоит из мембран и связанных с ними рибосом. Было бы интересно выяснить, существует ли в этой структуре прямое взаимодействие между липидами и РНК.

Краткие выводы

Исследование свойств некоторых мало изученных биологических полимеров и агрегатов весьма информативно и в то же время поднимает множество вопросов.

Полисахариды существуют в виде как линейных, так и разветвленных структур. Нередко они представляют собой гомополимеры или чередующиеся сополимеры. На уровне вторичной структуры они являются одно- и многоцепочечными спиралями. Далее эти спирали организованы в фибриллярные и гелеобразные (третичную и четвертичную) структуры. Сахариды являются также компонентами других биологических комплексов, таких, как пептидогликан клеточных стенок бактерий и гликопротеиды эукариотических клеток.

Комплексы между белками и нуклеиновыми кислотами выполняют очень важные биологические функции и по своей организации бывают весьма разнообразны. У простых ви-

русов основным компонентом, определяющим строение вирусной частицы, является белок, и он располагается снаружи структуры. В хроматине, напротив, белок образует сердцевину, на которую намотана ДНК. Примером структуры, в которой имеют место тесные взаимодействия между белками и нуклеиновыми кислотами, служат рибосомы. Детальные механизмы узнавания этими двумя типами макромолекул друг друга неизвестны.

Липиды — это амфифильные соединения; они образуют мицеллы, если содержат по одной жирнокислотной цепи, и двойные слои или бислойные пузырьки, если таких цепей две. Свойства и состав двух поверхностей бислоя не обязательно одинаковы. Природные мембраны помимо липидов содержат большое количество белков. Периферические белки легко экстрагируются из мембраны, в то время как интегральные мембранные белки прочно связаны с ней, вероятно, с помощью гидрофобного участка пептидной цепи. Некоторые интегральные цепи локализуются только на одной поверхности мембраны, другие пронизывают ее насквозь. В липидных бислоях происходят фазовые переходы между состояниями, которые условно можно считать твердым и жидким. В природных мембранах тоже наблюдаются аналогичные переходы, а также латеральное фазовое разделение. От других биологических структур мембраны отличает то, что они являются динамическими системами. В них происходит довольно быстрое латеральное перемещение белков и липидов и вращение различных компонентов. Однако перескок компонентов с одной поверхности на другую происходит весьма редко.

О деталях липидно-белковых взаимодействий известно совсем немного. Полезными модельными системами могут служить липопротеиды сыворотки, которые являются мицеллоподобными структурами, состоящими из разных пептидов и липидов. Эти соединения интересны и сами по себе, поскольку играют важную роль в транспорте липидов.

ЗАДАЧИ

4.1. Предположим, что вам удалось выделить мицеллу, содержащую одну молекулу белка, обычно существующего в виде трансмембранного белка. Каким, по вашему мнению, будет взаимное расположение в такой мицелле белка и липидов?

4.2. Для того чтобы идентифицировать белки, расположенные рядом друг с другом в клеточной мембране, часто используется метод сшивания. Однако из-за большой латеральной подвижности некоторые мембранные белки, первоначально расположенные далеко друг от друга, могут случайно оказаться рядом, что может привести к образованию между ними поперечных сшивок. Как бы вы выявили такие случаи или предотвратили их в эксперименте по сшиванию?

4.3. Почему есть основания ожидать, что олигомерный белок сможет участвовать в специфическом узнавании палиндромной последовательности нуклеиновой кислоты?

4.4. Почему нерегулярность последовательностей и ветвление цепи полисахаридов увеличивают вероятность образования геля, а не других возможных типов третичной и четвертичной структуры?

4.5. Одни нуклеопротеидные комплексы стабилизируются при увеличении концентрации солей, в то время как другие дестабилизируются. Чем могут различаться структуры этих двух типов комплексов? Рассмотрите типы белков, которые могут участвовать в образовании комплекса, возможные детали белково-нуклеиновых контактов и число малых катионов, связанных с комплексами, по сравнению с числом этих ионов в случае несвязанных компонентов.

ЛИТЕРАТУРА

Общая

Gennis R. B., Jonas A., 1977. Protein—lipid interactions, Ann. Rev. Biophys. Bioeng., **6**, 195.
Kirkwood S., 1974. Unusual polysaccharides, Ann. Rev. Biochem., **43**, 401.
Rogers H. J., Perkins H. R., 1968. Cell Walls and Membranes, London, Spon.
Singer S. J., 1974. The molecular organization of membranes, Ann. Rev. Biochem., **43**, 805.
Von Hippel P. H., McGhee J. D., 1972. DNA—protein interactions, Ann. Rev. Biochem., **41**, 231.

Специальная

Braun V., Hantke K., 1974. Biochemistry of bacterial cell envelopes, Ann. Rev. Biochem., **43**, 89.
Casjens S., King J., 1975. Virus assembly, Ann. Rev. Biochem., **44**, 555.
Champoux J. J., 1978. Proteins that affect DNA conformation, Ann. Rev. Biochem., **47**, 449.
DiRienzo J. M., Nakamura K., Inouye M., 1978. The outer membrane proteins of Gram-negative bacteria: Biosynthesis, assembly and functions, Ann. Rev. Biochem., **47**, 481.
Elgin S. C. R., Weintraub H., 1975. Chromosomal proteins and chromatin structure, Ann. Rev. Biochem., **44**, 900.
Kornfeld R., Kornfeld S., 1976. Comparative aspects of glycoprotein structure, Ann. Rev. Biochem., **45**, 217.
Lancelot G., Helene C., 1977. Selective recognition of nucleic acids by proteins: The specificity of guanine interactions with carboxylate ions, Proc. Natl. Acad. Sci. USA, **74**, 4872.
Marchesi V. T., Furthmayr H., Tomita M., 1976. The red cell membrane, Ann. Rev. Biochem., **45**, 667.
Morrisett J. D., Jackson R. L., Gotto A. M., Jr., 1975. Lipoproteins: Structure and function, Ann. Rev. Biochem., **44**, 183.
Rothman J. E., Lenard J., 1977. Membrane asymmetry, Science, **195**, 743.
Stanier R. Y., Adelberg F. A., Ingraham J. L., 1976. The Microbial World, 4th ed., Englewood Cliffs, N. J., Prentice-Hall. [Имеется перевод: Стейниер Р., Эдельберг Э., Ингрем Дж. Мир микробов. — М.: Мир, 1979. В гл. 11 подробно рассмотрено строение клеточных стенок.]
Tanford C., 1973. The Hydrophobic Effect, New York, Wiley.
Thomas J. O., 1978. Chromatin structure. In: Biochemistry of Nucleic Acids, vol. 2, ed. B. F. C. Clark, Baltimore, University Park Press, p. 181.
Warrant R. W., Kim S. H., 1978. α-Helix—double helix interaction shown in the structure of a protamine—transfer RNA complex and a nucleoprotamine model, Nature, **271**, 130.
Wittmann H. G., 1976. Structure, function and evolution of ribosomes, Eur. J. Biochem., **61**, 1.

Глава 5 Конформационный анализ и силы, определяющие структуру белка

5.1. Основные проблемы, связанные со структурой белковых молекул

Одна из наиболее трудных и интересных задач биохимии — это выяснение физических основ строения белковых молекул, отличающихся тонкой организацией и высокой специфичностью. Эта задача приобретает особое значение из-за того, что биологическая активность таких макромолекул чувствительна к изменениям их пространственной конформации. Так как биологические полимеры способны разворачиваться (например, при нагревании или обработке мочевиной) и затем снова сворачиваться, возвращаясь в исходное состояние, разумно предположить, что в основном конформации, принимаемые различными биополимерами, наиболее выгодны термодинамически. Некоторые данные, подтверждающие эту гипотезу, были получены в серии экспериментов, выполненных Анфинсеном и др. (1961 г.). Эти исследования показали, что даже при разрыве всех четырех дисульфидных связей рибонуклеазы и полной ее денатурации мочевиной после удаления мочевины и реокисления дисульфидных мостиков белок снова приобретает правильную нативную конформацию[1].

В этой главе мы обсудим различные факторы, определяющие конформацию белка. Сначала мы рассмотрим особенности геометрии молекулы, например фиксированные длины связей и величины валентных углов полипептидной цепи. Затем проанализируем ограничения, налагаемые на возможные конформации стерическими факторами. Мы также рассмотрим усовершенствованный вариант подобного анализа, основанный на использовании более реалистических потенциальных функций. В заключение мы познакомимся с другими факторами, играющими очень важную роль в формировании белковых структур. К ним относятся образование хорошо изученных водородных связей и «гидрофобные» взаимодействия, природа которых менее понятна. Если бы все эти факторы были достаточно хорошо изучены и существовал соответствующий математический аппарат, появилась бы возможность предсказывать, например, трехмерную структуру белка по его аминокислотной последовательности. Эта цель еще не достигнута, но недавние успехи показывают, что ее нельзя считать нереальной. (Другие аспекты укладки белковых молекул обсуждаются в гл. 21.)

5.2. Геометрия полипептидной цепи

На рис. 5.1 представлена α-L-полипептидная цепь, все остатки которой находятся в *транс*-положении (так называемый плоский зигзаг). Остатки пронумерованы последовательно от 1 до n, начиная с N-конца. Отметим, что одинаковые атомы одного и того же

[1] Несмотря на результаты этих экспериментов, необходимо считаться с возможностью того, что нативным структурам соответствует локальный минимум свободной энергии, а достижению конформации с более низкой свободной энергией препятствуют кинетические барьеры.

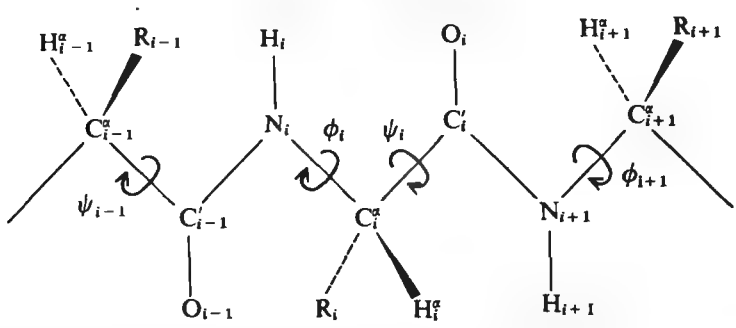


РИС. 5.1. α-L-полипептидная цепь. Все аминокислотные остатки находятся в *транс*-конформации.

остатка обозначаются по-разному: например, *i*-й α-углерод обозначается C^{α}_i, тогда как *i*-й карбонильный углерод — C'_i. Индивидуальные аминокислотные остатки различаются своими боковыми цепями (часто называемыми R-группами). Длины связей и валентные углы, определяющие скелет цепи, можно считать фиксированными. В табл. 5.1 приведены общепринятые значения этих параметров.

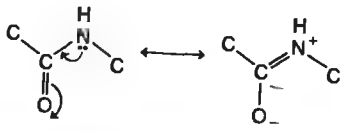
Таблица 5.1

ГЕОМЕТРИЯ ПОЛИПЕПТИДНОЙ ЦЕПИ[1)]

Связь	Длина связи, Å	Связи	Валентный угол, град.
$C^{\alpha}—C'$	1,53	$C^{\alpha}—C'—N$	113
$C'—N$	1,32	$C'—N—C^{\alpha}$	123
$N—C^{\alpha}$	1,47	$N—C^{\alpha}—C'$	110
$C'=O$	1,24		
$N—H$	1.00		
$C^{\alpha}—C^{\beta}$	1,54		
$C^{\alpha}—H^{\alpha}$	1,07		

[1]. Sasisekharan. In: Collagen, ed. N. Ramanathan (New York: Interscience, 1962), p. 39.

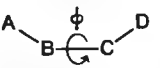
Важная особенность структуры состоит в том, что амидная группа обычно находится в плоской *транс*-конформации. Это можно объяснить резонансным состоянием, придающим амидной связи характер двойной связи,



в результате чего эта связь укорачивается на 0,1 Å. Вследствие плоской *транс*-конформации амидной группы расстояние между последовательными α-углеродными атомами постоянно и равно 3,8 Å. Отсюда следует, что конформация полипептидной цепи как целого определяется углами поворота вокруг связей $N—C^{\alpha}$ и $C^{\alpha}—C'$.

УГЛЫ ВНУТРЕННЕГО ВРАЩЕНИЯ

Для описания трехмерной структуры макромолекулы необходимо определить внутренние углы вращения, или так называемые торсионные углы. Мы следуем здесь рекомендациям IUPAC — IUB (1970 г.). Рассмотрим систему, состоящую из четырех атомов A—B—C—D.

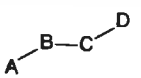


При заданных длинах связей и валентных углах торсионный угол ϕ описывает взаимное расположение связей C—D и A—B. Он определяется как угол между плоскостью, задаваемой связями B—C и C—D, и плоскостью, задаваемой связями A—B и B—C.

Численные значения, приписываемые по соглашению торсионному углу, определяются следующим образом. На рис. 5.2, *А*,*Б* показаны проекции Ньюмена и проекции рассматриваемой нами четырехатомной системы в перспективе. Положительное значение угла ϕ присваивается правому (происходящему по часовой стрелке) повороту, необходимому для перевода ближнего атома (A или D) в «заслоненное» положение отиосительно дальнего (D или A) (когда ближний атом заслоняет дальний на проекции Ньюмена). Например, на первой и третьей проекциях рис. 5.2,*А* правый поворот ближнего к нам атома A на угол ϕ_1 вокруг связи B—C (если смотреть в направлени от B к C) приводит к тому, что атом A заслоняет атом D. Если атомы расположены в обратном порядке, как на второй и четвертой проекциях, то правый поворот атома D на угол ϕ_2 вокруг связи C—B также приводит к заслоненному положению атомов D и A. Таким образом, независимо от того, смотрим ли мы в направлении от B к C или от C к B, мы получаем то же положительное значение для угла ϕ.

С другой стороны, это же значение ϕ получается, если мы смотрим в направлении от B к C и определяем ϕ как правый поворот связи C—D (вокруг B—C), необходимый для перевода атомов A (ближний) и D (дальний) из заслоненного положения в исходное. Очевидно, мы можем выполнить эквивалентную операцию и глядя в направлении от C к B. Отрицательные значения присваиваются торсионному углу при левом (происходящем против часовой стрелки) повороте (рис. 5.2, *Б*).

Значение $\phi = 0°$ соответствует заслоненной конформации, в которой связи A—B и C—D находятся в *цис*-положении. Когда они находятся в *транс*-положении



торсионный угол $\phi = 180°$ (или, что то же, $\phi = -180°$). Удобно определять значение ϕ в области $-180° \leqslant \phi \leqslant +180°$; тогда величина этого угла показывает, какая связь существует между энантиомерными конформациями.

Для описания структуры скелета полипептидной цепи с фиксированными в *транс*-конформации амидными группами важными торсионными углами являются углы поворота относительно связей $N_i—C_i^\alpha$ и $C_i^\alpha—C_i'$. Эти углы обозначаются соответственно через ϕ_i и ψ_i. Для α-L-полипептида в плоской *транс*-конформации (рис. 5.2, *В*) $\phi_i = \psi_i = 180°$ для всех i.

[Следует отметить, что опубликованная ранее литература по этому вопросу основывалась на соглашении, по которому исходному состоянию соответствует $\phi_i = 0°$ и $\psi_i = 0°$ при отсчете углов от плоской *транс*-конформации. Поэтому при чтении такой литературы следует к числовым значениям ϕ и ψ добавлять (или вычитать из них) 180°, для того чтобы они соответствовали правилу, принятому в даниой книге. Более детальное обсуждение

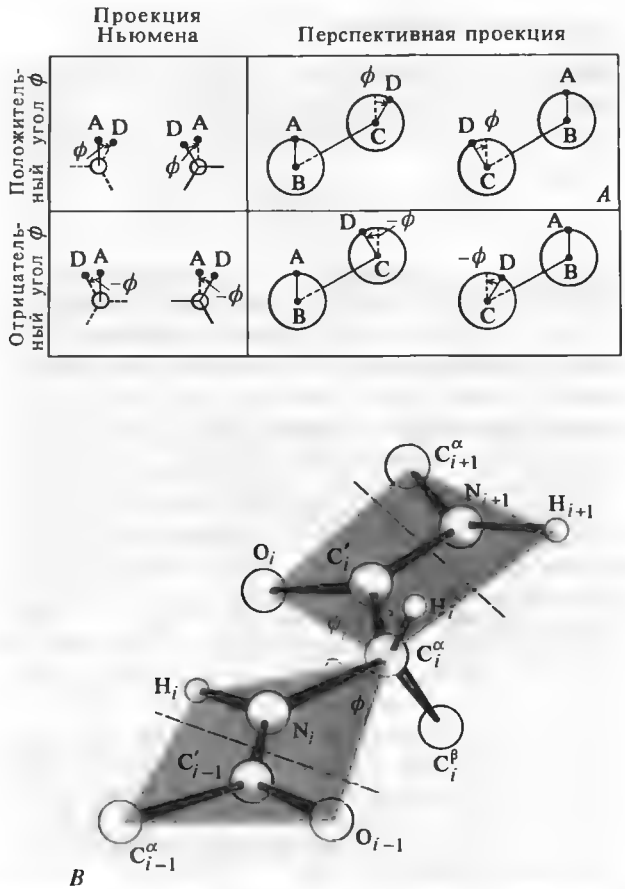


РИС. 5.2. Торсионные углы. *А*. Проекции Ньюмена и перспективные проекции, иллюстрирующие ситуацию, в которой торсионный угол ϕ для системы из четырех атомов (A, B, C, D) имеет положительное значение. *Б*. Аналогичные проекции, иллюстрирующие отрицательные значения угла ϕ. *В*. Проекция участка α-L-полипептидной цепи с указанием углов ϕ и ψ. Цепь представлена в форме плоского зигзага (остатки находятся в *транс*-конформации), для которого $\phi = \psi = 180°$. Пунктирные линии очерчивают i-й аминокислотный остаток. [IUPAC-IUB Commission on Biochemical Nomenclature, Biochemistry, 9, 3471 (1979).]

структурных соглашений, особенно тех, которые используются при рассмотрении статистики цолимерных цепей, можно найти в Дополнении 18.1.]

Для определения конформации скелета полипептидной цепи как целого необходимо указать значения всех ϕ_i и ψ_i. Если $\phi_i = \phi_j$ и $\psi_i = \psi_j$, полипептид имеет спиральную конформацию. Для правой α-спирали $\phi = -57°$ и $\psi = -47°$. Данные о φ и ψ для целого ряда упорядочных конформаций полипептидов приведены в табл. 5.2. Все они являются периодическими структурами, в которых углы поворота для каждого из остатков одинаковы. (Некоторые такие структуры показаны в гл. 2). Отметим, что конформация складчатого слоя наиболее близка к плоской зигзагообразной структуре, показанной на рис. 5.1.

В отличие от упорядоченных полипептидов, перечисленных в табл. 5.2, величины ϕ, ψ в глобулярных белках (и, конечно, в случайно свернутых клубках) значительно варьируют от остатка к остатку.

Таблица 5.2

ПРИБЛИЗИТЕЛЬНЫЕ ЗНАЧЕНИЯ ТОРСИОННЫХ УГЛОВ ДЛЯ НЕКОТОРЫХ РЕГУЛЯРНЫХ СТРУКТУР[1]

Структура	ϕ, град.	ψ, град.
Правая α-спираль [α-поли(L-аланин)]	−57	−47
Левая α-спираль	+57	+47
β-Слой с параллельными цепями	−119	+113
β-Слой с антипараллельными цепями [β-поли(L-аланин)]	−139	+135
Полиглицин II	−80	+150
Коллаген (тройная спираль)	−51, −76, −45	+153, +127, +148
Поли(L-пролин) I	−83	+158
Поли(L-пролин) II	−78	+149

[1] IUPAC-IUB Comission on Biochemical Nomenclature, Biochemistry, **9**, 3471 (1970); одновременная публикация в J. Mol. Biol., **52**, 1(1970) и J. Biol. Chem., **245**, 6489 (1970).
У полностью вытянутой цепи $\phi = \psi = +180°$. Имидные связи полипролина I находятся в *цис*-, полипролина II — в *транс*-конформациях.

КОНФОРМАЦИОННЫЕ КОНТУРНЫЕ КАРТЫ, ИЛИ ДИАГРАММЫ РАМАЧАНДРАНА

Теоретически скелет линейной молекулы может иметь бесконечное множество конформаций, каждой из которых соответствует единственный набор значений углов поворота. Однако многие из этих конформаций можно исключить из-за стерических факторов. Г. Рамачандран и его сотр. одними из первых исследовали эту проблему.

При анализе пространственных взаимодействий прежде всего необходимо рассмотреть наложение атомов, расстояния между которыми зависят от одного угла поворота. Например, расстояния между атомами C_i^α и O_i, C_i^α и N_{i+1} или C_i^α и C_{i+1}^α зависят только от ψ_i[1]. Те углы ψ_i, при которых возникают стерические затруднения, необходимо исключить из числа возможных. Однако при изучении моделей в атомном масштабе становится очевидно, что можно исключить еще некоторые значения угла ψ_i. Это объясняется наложением атомов, расстояния между которыми зависят от двух углов, ϕ_i и ψ_i, скажем атомов O_i и O_{i-1}. Из-за таких наложений повороты вокруг связей $N_i—C_i^\alpha$ и $C_i^\alpha—C_i'$ являются взаимозависимыми.

Рассматривая взаимодействия между всеми атомами, расстояния между которыми зависят только от этих двух углов, можно определить стерически разрешенные значения пары ϕ_i и ψ_i. Области разрешенных значений ϕ_i и ψ_i *не* сокращаются при учете взаимодействий, зависящих одновременно от поворотных углов соседних остатков, таких, как ϕ_{i-1} и ψ_{i-1}. Например, расстояние между O_i и R_{i-1} зависит от углов ϕ_i, ϕ_i и ϕ_{i-1}, но когда углы ϕ_i и ψ_i находятся в пределах ранее определенных допустимых областей, никакие контакты между O_i и R_{i-1} невозможны. Таким образом, при фиксированной в *транс*-положении амидной группе повороты в данном остатке взаимозависимы, но в соседних остатках стерически независимы.

[1] Расстояния между этими атомами не зависят от каких-либо углов поворота, так как они принадлежат пептидной группе, которую мы условились считать фиксированной в плоской *транс*-форме. — *Прим. перев.*

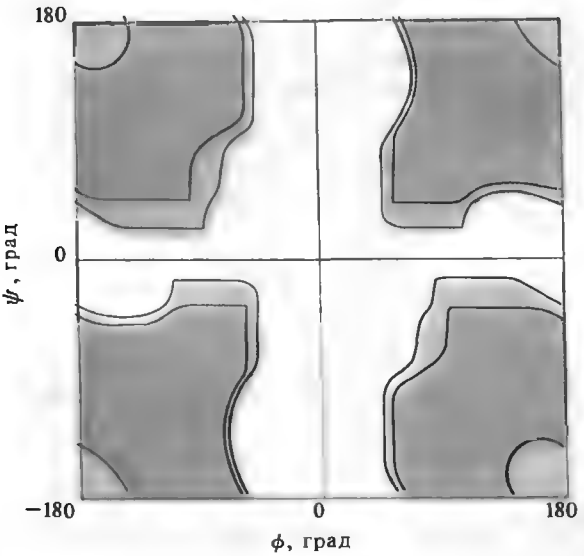


РИС. 5.3. Коиформациоииая контурная карта для глицила в полипептидной цепи. Темные и светлые зоиы — области обычиых и предельных стерически допустимых коитактов соответственно. [G. Ramachandran et al., Biophys. J., **6**, 849 (1966).]

Здесь важно провести различие между ближними и дальними взаимодействиями. Взаимодействия атомов и групп, соседних или расположенных недалеко друг от друга вдоль белковой цепи, называются ближними. Именно они обсуждались выше. Взаимодействия между группами, расположенными в разных местах цепи, называются дальними, они происходят только тогда, когда при складывании молекулы такие группы сближаются в пространстве. Данный параграф посвящен в основном чрезвычайно важным для понимания конформаций белков ближним взаимодействиям.

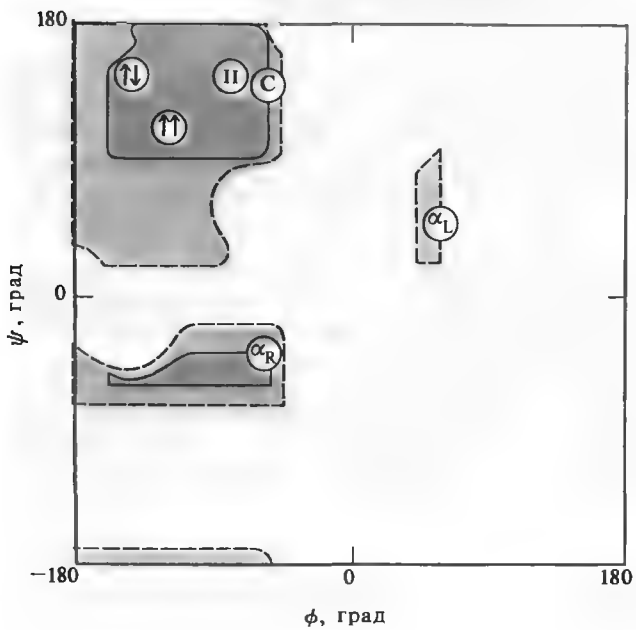


РИС. 5.4. Конформационная коитурная карта для L-аланила в полипептидиой цепи. Темиые и светлые зоны — области обычных и предельных стерически допустимых контактов соответственно. Отмечены координаты правой и левой α-спиралей (α), β-слоев с параллельными (↑↑) и антипараллельными (↑↓) цепями, полиглицина II (II) и коллагена (C) (см. также табл. 5.2). [P.J. Flory, Statistical Mechanics of Chain Molecules (New York: Interscience, 1969); G.N. Ramachandran et al., J.Mol.Biol., **7**, 95 (1963).]

Таблица 5.3

НАИМЕНЬШИЕ МЕЖАТОМНЫЕ РАССТОЯНИЯ ДЛЯ ВАЖНЫХ ПАР АТОМОВ В ПОЛИПЕПТИДАХ[1]

Пары атомов	Обычное расстояние, Å	Предельно допустимое расстояние, Å	Пары атомов	Обычное расстояние, Å	Предельно допустимое расстояние, Å
C—C	3,2	3,0	O—N	2,7	2,6
C—O	2,8	2,7	O—H	2,4	2,2
C—N	2,9	2,8	N—N	2,7	2,6
C—H	2,4	2,2	N—H	2,4	2,2
O—O	2,8	2,7	H—H	2,0	1,9

[1] Расстояния O—O по C. Ramakrishnan, G. N. Ramachandran, Biophys. J., 5, 909(1965); остальные данные по G. N. Ramachandran et al., J. Mol. Biol., 7, 95(1963).

Как наилучшим образом графически представить области разрешенных значений углов ϕ_i и ψ_i для данного аминокислотного остатка? Это делается при помощи карты, на которой по оси ординат отложен угол ψ, а по оси абсцисс — угол ϕ. На этих картах линиями выделяются области значений ϕ, ψ, в которых отсутствуют стерические ограничения.

На рис. 5.3 представлена стерическая карта для глицила (R = H), расположенного в полипептидной цепи, на рис. 5.4 — аналогичная карта для аланила (R = CH_3). Темные области соответствуют разрешенным значениям углов ϕ и ψ при обычных межатомных расстояниях, более светлые — при предельных межатомных расстояниях, наблюдаемых в кристаллических структурах молекул. Эти предельные межатомные расстояния определяют «внешнюю границу» стерически допустимых конформаций.

В табл. 5.3 приведены наименьшие межатомные расстояния для различных пар атомов, использованные в рис. 5.3. Эти данные основаны на значениях, найденных при определении структуры методом рентгеноструктурного анализа.

Конформационная карта для глицина центросимметрична, т.е. симметрична относительно любой линии, проходящей через ее центр ($\phi = 0°$, $\psi = 0°$). Это есть следствие симметрии самого глицинового остатка. Как можно было ожидать, *цис*-конфигурации относительно любой из двух связей ($\phi = 0°$ или $\psi = 0°$) и конфигурации, близкие к ним, стерически запрещены. В случае аланина стерическая карта асимметрична и стерически разрешенная область значительно меньше по размеру. Это связано с наличием дополнительных стерических ограничений, налагаемых присутствием группы — CH_3 боковой цепи. Конформационные карты для остатков с более длинной, не разветвляющейся при β-углеродном атоме боковой цепью, как, например, лейцин, R = $-CH_2CH(CH_3)_2$ —, весьма похожи на рис. 5.4. Для остатков с разветвляющейся при β-углеродном атоме боковой цепью, таких, как валин (R = $CH(CH_3)_2$), стерически разрешенные области меньше областей на рис. 5.4.

5.3. Оценки потенциальной энергии

Конформационные карты на рис. 5.3 и 5.4 дают приближенное представление о разрешенных конформациях аминокислотных остатков. Они, однако, не указывают, какие части стерически разрешенных областей наиболее предпочтительны, т.е. они не дают внутри стерически разрешенных областей градации по энергии. Для того чтобы оценить относительную предпочтительность отдельных конформаций, необходимо найти зависимость потенциальной энергии от ϕ и ψ.

Существующие методы расчета потенциальной энергии как функции углов поворота — полуэмпирические, поэтому к значениям энергии, полученным таким образом, нельзя относиться слишком доверчиво. Однако при использовании этих методов в сочетании с конфигурационной статистикой (см. гл. 18) обычно наблюдается разумное согласие между рассчитанными и экспериментально найденными размерами полимерных цепей. Вместе с тем использование карт, полученных методом «твердых сфер» (как на рис. 5.3 и 5.4), не всегда дает хорошее согласие с экспериментом.

ВЗАИМОДЕЙСТВИЯ МЕЖДУ ВАЛЕНТНО НЕ СВЯЗАННЫМИ АТОМАМИ

Рассмотрим вначале отталкивание и притяжение атомов, расстояние между которыми зависит от ϕ и ψ. Эти взаимодействия между валентно не связанными атомами (невалентные взаимодействия) и дают энергию $E_{kl}(\phi_i, \psi_i)$, которая обычно записывается в виде

$$E_{kl}(\phi_i, \psi_i) = a_{kl}/r_{kl}^m - c_{kl}/r_{kl}^6 \tag{5.1}$$

где параметры a_{kl} и c_{kl} характеризуют пару k- и l-атомов или атомных групп (например, CH_3). Типичная зависимость E_{kl} от r_{kl} приведена на рис. 5.5. График показывает, что на больших расстояниях взаимодействие отсутствует. По мере сближения атомов растет притяжение между ними, а на расстояниях, на которых атомные радиусы перекрываются, начинает преобладать сильное отталкивание.

Член выражения (5.1), характеризующий притяжение и зависящий от расстояния между атомами в степени -6, есть не что иное, как хорошо известная дисперсионная энергия Лондона. Распределение электронной плотности вокруг атомного ядра испытывает частые флуктуации, что порождает смещение центров электронных и ядерных зарядов. Эти смещения приводят к флуктуациям дипольного момента, которые могут находиться в резонансе с аналогичными флуктуациями дипольного момента соседнего атома, что приводит к их постоянному притяжению. Это взаимодействие индуцированный диполь — индуцированный диполь и объясняет зависимость энергии от шестой степени расстояния. Параметр c_{kl} получается из величин поляризуемости атомов k и l и из эффективного числа электронов, способных поляризоваться. (См. Brant et al., 1967 и приведенную там библиографию.)

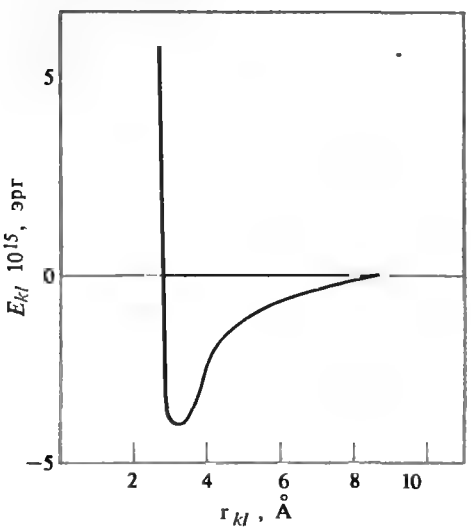


РИС. 5.5. Зависимость энергии невалентных взаимодействий E_{kl} от расстояния между взаимодействующими атомами или группами с изомерами k и l (в данном случае — между двумя молекулами водорода). [W.J. Moore, Physical Chemistry, 3d ed. (Englewood Cliffs, N.Y.: Prentice-Hall, 1962), p. 715].

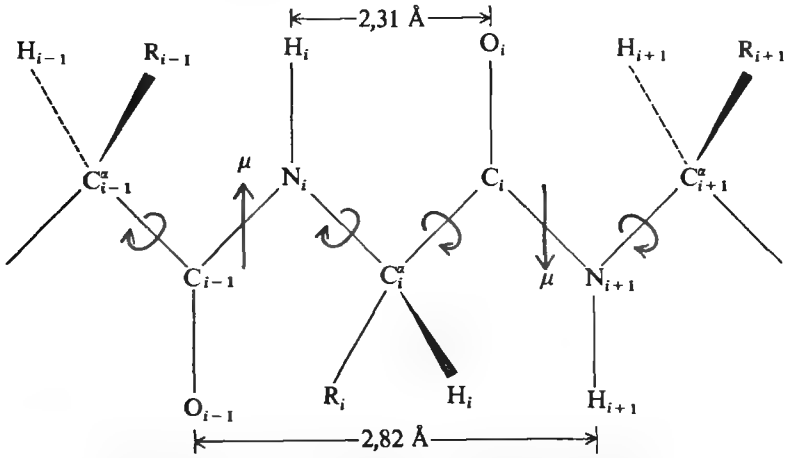


РИС. 5.6. Схема, иллюстрирующая наличие дипольного момента у амидной группы. В вытянутой цепи направления моментов чередуются. [P.J. Flory, Statistical Mechanics of Chain Molecules (New York: Interscience, 1969), chap. VII.]

Член a_{kl}/r_{kl}^m в выражении (5.1), описывающий энергию отталкивания, не имеет строгого теоретического обоснования[1]. Результаты, полученные в экспериментах с молекулярными пучками, показывают, что m варьирует от 9 до 12. Параметр a_{kl} можно получить из требования, чтобы минимум кривой на рис. 5.5 соответствовал известным межатомным расстояниям (см. также Brant et al., 1967). При расчетах структуры полипептидов обычно используют следующие величины атомных радиусов: $r_H^0 = 1{,}2 \div 1{,}3$ Å; $r_O^0 = 1{,}5 \div 1{,}6$ Å; $r_{CH_2} = 1{,}95$ Å (CH_2-группа обычно рассматривается как единая сфера).

ДИПОЛЬНЫЕ ВЗАИМОДЕЙСТВИЯ

Другой вклад в энергию, зависящую от поворотов на углы ϕ и ψ, дают дипольные взаимодействия между соседними амидными группами. Дипольные моменты амидных групп приблизительно параллельны N—H-связи и направлены от N к H (рис. 5.6). Их численное значение примерно равно 3,7 Д. Это довольно большой момент; например, дипольный момент HCl равен 1,03 Д, CH_3Cl — 1,87 Д и HCN — 2,93 Д. Поэтому мы вправе ожидать, что дипольные взаимодействия оказывают значительное влияние на конформацию скелета полипептидной цепи.

Ориентация дипольных моментов любых двух последовательно расположенных амидных групп определяется, несомненно, углами φ и ψ. Энергия E_d взаимодействия двух точечных диполей с моментами μ_A и μ_B, находящихся на расстоянии r, определяется формулой

$$E_d = \varepsilon^{-1}[\mu_A \cdot \mu_B/r^3 - 3(\mu_A \cdot \mathbf{r})(\mu_B \cdot \mathbf{r})/r^5] \qquad (5.2)$$

где r — модуль вектора r, а ε — диэлектрическая проницаемость (см. Дополнение 5.1). Это выражение, конечно, весьма приближенно, поскольку оно справедливо только для расстояний между амидными диполями, достаточно больших по сравнению с расстояниями между парциальными зарядами в самих диполях. Для более точного представления об электростатических взаимодействиях между соседними амидными диполями можно, приписав парциальные заряды соответствующим атомам, рассчитать полную электро-

[1] Можно использовать также экспоненциальный потенциал отталкивания. (См., например, Brant, Flory, 1965.)

Дополнение 5.1

ВЫВОД ВЫРАЖЕНИЯ ДЛЯ ЭНЕРГИИ ВЗАИМОДЕЙСТВИЯ ДВУХ ДИПОЛЕЙ

Энергию взаимодействия двух точечных диполей не очень трудно рассчитать из общих соображений. Если обозначить соответственные дипольные моменты через μ_A и μ_B, то энергия их взаимодействия просто равна потенциальной энергии μ_B в электрическом поле, образованном μ_A (или наоборот), т.е. равна $-\mathbf{E}_A\mu_B$, где $\mathbf{E}_A$ — напряженность электрического поля диполя с моментом μ_A. Таким образом, мы сводим задачу к расчету $\mathbf{E}_A$.

Прежде всего рассмотрим общий случай поля, созданного произвольно распределенными зарядами, и проиллюстрируем концепцию мультипольного разложения. Пусть в некоторой системе координат i-й заряд $z_i e$ имеет координаты $\tilde{x}_i$, $\tilde{y}_i$, $\tilde{z}_i$. Потенциал электрического поля на некотором расстоянии $\mathbf{r} = \hat{\mathbf{i}}\tilde{x} + \hat{\mathbf{j}}\tilde{y} + \hat{\mathbf{k}}\tilde{z}$ от начала координат определяется формулой

$$\Phi = \sum [z_i e/\varepsilon][(\tilde{x} - \tilde{x}_i)^2 + (\tilde{y} - \tilde{y}_i)^2 + (\tilde{z} - \tilde{z}_i)^2]^{1/2}]$$

Это выражение можно разложить в ряд Тейлора в точке $\tilde{x}_i$, $\tilde{y}_i$, $\tilde{z}_i$. В результате получим

$$\Phi = (1/\varepsilon)\{\sum (z_i e/r) + \sum [z_i e\tilde{x}_i(\partial X^{-1/2}/\partial\tilde{x}_i)] + \sum [z_i e\tilde{y}_i(\partial X^{-1/2}/\partial\tilde{y}_i)] + \sum [z_i e\tilde{z}_i(\partial X^{-1/2}/\partial\tilde{z}_i)] + \\ + \frac{1}{2}\sum [z_i e\tilde{x}_i^2(\partial^2 X^{-1/2}/\partial\tilde{x}_i^2)] + \sum [z_i e\tilde{x}_i\tilde{y}_i(\partial^2 X^{-1/2}/\partial\tilde{x}_i\partial\tilde{y}_i)] + \ldots]$$

где $x = (\tilde{x} - \tilde{x}_i)^2 + (\tilde{y} - \tilde{y}_i)^2 + (\tilde{z} - \tilde{z}_i)^2$, а все производные берутся в точке $\tilde{x}_i = \tilde{y}_i = \tilde{z}_i = 0$.

Отметим, что первый член есть просто потенциал электростатического поля в точке r, если все заряды $z_i e$ сконцентрированы в начале координат [потенциал = (суммарный заряд)/εr]. Следующие члены вносят поправки, отражающие то обстоятельство, что заряды как-то распределены вокруг начала координат. Первые три из них содержат компоненты дипольного момента μ:

$$\mu_{\tilde{x}} = \sum z_i e\tilde{x}_i; \quad \mu_{\tilde{y}} = \sum z_i e\tilde{y}_i; \quad \mu_{\tilde{z}} = \sum z_i e\tilde{z}_i$$

Девять следующих членов содержат девять компонент квадрупольного момента $\mathbf{Q}$ распределения зарядов:

$$Q_{\tilde{x}\tilde{x}} = \sum z_i e\tilde{x}_i^2; \quad Q_{\tilde{y}\tilde{y}} = \sum z_i e\tilde{y}_i^2; \quad Q_{\tilde{z}\tilde{z}} = \sum z_i e\tilde{z}_i^2; \quad Q_{\tilde{y}\tilde{x}} = \sum z_i e\tilde{x}_i\tilde{y}_i; \ldots$$

Эти девять компонент образуют тензор квадрупольного момента.

Используя связи между компонентами дипольного и квадрупольного моментов и замечая, что $(\partial X^{-1/2}/\partial\tilde{x}_i)_{\tilde{x}_i=0} = -\partial r^{-1}/\partial\tilde{x}$ (аналогично для всех производных по $\partial\tilde{y}_i$ и $\partial\tilde{z}_i$), можно записать разложение Φ в ряд следующим образом:

$$\Phi = (1/\varepsilon)[(\sum z_i e)/r + \mu_{\tilde{x}}(\partial r^{-1}/\partial\tilde{x}) - \mu_{\tilde{y}}(\partial r^{-1}/\partial\tilde{y}) - \mu_{\tilde{z}}(\partial r^{-1}/\partial\tilde{z}) + \frac{1}{2}Q_{\tilde{x}\tilde{x}}(\partial^2 r^{-1}/\partial\tilde{x}^2) + \\ + \frac{1}{2}Q_{\tilde{y}\tilde{y}}(\partial^2 r^{-1}/\partial\tilde{y}^2) + \frac{1}{2}Q_{\tilde{z}\tilde{z}}(\partial^2 r^{-1}/\partial\tilde{z}^2) + Q_{\tilde{x}\tilde{y}}(\partial^2 r^{-1}/\partial\tilde{x}\partial\tilde{y}) + \ldots]$$

Это называется мультипольным разложением потенциала Φ; члены, содержащие компоненты дипольного момента, образуют дипольный потенциал, члены, содержащие вторые производные, — квадрупольный потенциал.

Предположим теперь, что существует одинаковое число зарядов разного знака (так что $\sum z_i e = 0$), расположенных на небольших расстояниях от начала координат и обладающих дипольным моментом μ_A (например, амидный диполь эквивалентен двум равным единичным зарядам разного знака, расположенным на расстоянии 3,7 Å друг от друга). Основным значащим членом в разложении Φ является дипольный потенциал, который (для случая небольших смещений зарядов) преобладает над квадрупольным потенциалом, зависящим от квадрата смещения. Отметив, что $\partial r^{-1}/\partial\tilde{x} = -\tilde{x}/r^3$, получим

$$\Phi = (\mu_{A\tilde{x}}\tilde{x} + \mu_{A\tilde{y}}\tilde{y} + \mu_{A\tilde{z}}\tilde{z})/\varepsilon r^3$$

где $\mu_{A\tilde{x}}$, $\mu_{A\tilde{y}}$ и $\mu_{A\tilde{z}}$ — компоненты μ_A. Напряженность электрического поля $\mathbf{E}_A$ равна

$$\mathbf{E}_A = -\nabla\Phi = -\nabla\mu_A \cdot \mathbf{r}/\varepsilon r^3 = -\mu_A/\varepsilon r^3 - (\mu_A \cdot \mathbf{r}/\varepsilon)\nabla(1/\mathbf{r}^3)$$

Потенциальная энергия взаимодействия с диполем $\mu_B = \hat{\mathbf{i}}\mu_{B\tilde{x}} + \hat{\mathbf{j}}\mu_{B\tilde{y}} + \hat{\mathbf{k}}\mu_{B\tilde{z}}$, расположенным от μ_A на расстоянии $\mathbf{r}$, равна $-\mathbf{E}_A\mu_B$, что после упрощения можно записать следующим образом:

$$E_d = (\mu_A \cdot \mu_B/\varepsilon r^3) - 3(\mu_A \cdot \mathbf{r})(\mu_B \cdot \mathbf{r})/\varepsilon r^5$$

Этот результат совпадает с уравнением (5.2).

статическую энергию как сумму энергий кулоновских взаимодействий. В этом случае энергия рассчитывается по формуле

$$E_{\text{д}} = \sum_{ij} \left(q_i q_j / \varepsilon r_{ij} \right) \tag{5.3}$$

где q_i и q_j — парциальные заряды на атомах i и j, а r_{ij} — расстояние между ними. Парциальные заряды должны быть выбраны таким образом, чтобы они приблизительно соответствовали как общему дипольному моменту амидной группы, так и принятым значениям для моментов связей N—H и C—O. Этому критерию удовлетворяют значения парциальных зарядов $-0{,}28e$ и $+0{,}28e$ для атомов N и H и $-0{,}39e$ и $+0{,}39e$ для атомов O и C соответственно.

Выбор значения диэлектрической проницаемости — непростая задача. Хотя макроскопическая диэлектрическая проницаемость воды приблизительно равна 80, локальная микроскопическая проницаемость должна быть гораздо меньше. Можно ожидать, что силовые линии между двумя *соседними* диполями проходят в значительной своей части через скелет самого полимера, а не через окружающую его воду. Поэтому более подходящей мерой локальной диэлектрической проницаемости вдоль полипептидного скелета в растворе может служить диэлектрическая проницаемость при высоких частотах твердых амидов и полиамидов, вклад в которую, зависящий от ориентации постоянных диполей, весьма мал. Соответствующие экспериментальные данные показывают, что значения ε, лежащие в пределах 2÷5, являются реальной оценкой эффективной диэлектрической проницаемости в области между двумя последовательно расположенными амидными группами в полипептидной цепи.

С хорошим приближением взаимодействиями дипольных групп, разделенных одним аминокислотным остатком, можно пренебречь, так как энергия дипольного взаимодействия быстро падает с увеличением расстояния между диполями [см. (5.2)]. Более того, на больших расстояниях силовые линии электростатического поля могут проходить через растворитель, что увеличивает эффективную диэлектрическую проницаемость.

СОБСТВЕННЫЕ ТОРСИОННЫЕ ПОТЕНЦИАЛЫ, ПРИСУЩИЕ ОДИНАРНЫМ СВЯЗЯМ

Помимо факторов, рассмотренных выше, существует также небольшое торможение вращения, обусловленное торсионным потенциалом, свойственным одинарным связям. Кроме невалентных и электростатических взаимодействий каждая связь имеет собственный потенциальный барьер поворота. Поэтому повороты на углы ϕ и ψ порождают каждый свой, присущий ему, по всей видимости «тройной» (т.е. с тремя минимумами), потенциал заторможенного вращения. Эти минимумы приходятся на углы поворота, равные примерно 60, 180 и 300°. Полагают, что высоты барьеров между минимумами невелики, порядка 1 ккал/моль. Таким образом, торсионную энергию $E_{\text{тор}}(\phi_i, \psi_i)$ можно представить в виде

$$E_{\text{тор}}(\phi_i, \psi_i) = (E_\phi^0/2)(1 + \cos 3\phi) + (E_\psi^0/2)(1 + \cos 3\psi) \tag{5.4}$$

где E_ϕ^0 и E_ψ^0 — высоты барьеров для поворотов на углы ϕ и ψ соответственно.

ОБЩАЯ ЭНЕРГИЯ КАК СУММА ОТДЕЛЬНЫХ ВКЛАДОВ

Рассчитаем теперь полную потенциальную энергию как функцию углов ϕ_i и ψ_i для i-го аминокислотного остатка, расположенного внутри полипептидной цепи. Для этого необходимо рассчитать, как меняются расстояния между всеми атомами по мере изменения угла ϕ_i и/или ψ_i. Для каждой пары атомов k и l энергия невалентных взаимодействий E_{kl} вычисляется согласно уравнению (5.1), и сумма по всем атомным парам дает полную энергию невалентных взаимодействий. Расстояние между двумя диполями, расположенными по обе стороны от i-го α-углеродного атома, зависит от углов ϕ_i и ψ_i, и поэтому

энергия дипольных взаимодействий должна быть рассчитана для всех значений этих углов [уравнение (5.2) или (5.3)]. И наконец, для всех ϕ_i и ψ_i должен быть рассчитан вклад торсионного потенциала [уравнение (5.4)]. Таким образом, полный потенциал можно представить в виде

$$E(\phi_i, \psi_i) = \sum_{kl} [E_{kl}(\phi_i, \psi_i) + E_{\text{d}}(\phi_i, \psi_i) + E_{\text{тор}}(\phi_i, \psi_i)] \qquad (5.5)$$

где суммирование проводится по всем парам атомов, расстояния между которыми зависят от ϕ_i и/или ψ_i. Уравнение (5.5) служит основой для расчета конформационной энергии полипептидных цепей [1].

5.4. Результаты расчетов потенциальной энергии

ОСТАТКИ ГЛИЦИЛ И L-АЛАНИЛ

Результаты энергетических оценок, полученных согласно уравнению (5.5), показаны на рис. 5.7 для глицила и на рис. 5.8 для аланила. В обоих случаях остатки расположены внутри полипептидной цепи. Контуры энергетических зон проведены с интервалом в 1 ккал/моль относительно минимума, отмеченного крестиком. У глицила (рис. 5.7) существуют два симметрично расположенных минимума приблизительно при $\phi = -80°$, $\psi = +90°$ и $\phi = +80°$, $\psi = -90°$ и четыре области, в которых энергия остатка меньше 5 ккал/моль. Эти области близки к стерически разрешенным зонам, показанным на рис. 5.3.

Три различные низкоэнергетические области для L-аланила пронумерованы римскими цифрами (рис. 5.8). Точки, соответствующие наиболее низкой энергии в областях I и II, примерно отвечают координатам правой и левой α-спиралей соответственно (см. табл. 5.2)[2] Очевидно, однако, что минимум области I ниже, чем минимум области II. Следовательно, энергетическая карта дает разумное обоснование хорошо известному предпочтительному образованию L-полипептидами правых α-спиральных конформаций.

РОЛЬ ДИПОЛЬНЫХ ВЗАИМОДЕЙСТВИЙ

Область III является наиболее низкоэнергетической частью карты, представленной на рис. 5.8, причем зона с энергией 2 ккал занимает относительно большую площадь вокруг минимума (отмечен крестиком). Этой области соответствуют структуры с вытянутой конформацией, такие как хорошо известная β-структура (см. табл. 5.2). Следует, однако, отметить, что с энергетической точки зрения предпочтение остатками области III перед областью I обусловлено дипольными взаимодействиями амидных групп. Если эти взаимодействия не учитывать, область I становится предпочтительнее области III. Необходимость включать в расчеты дипольные взаимодействия получила экспериментальное подтверждение в опытах по измерению размеров статистических клубков цепей, подобных

---

[1] Это уравнение не учитывает взаимодействий с растворителем, которые также зависят от конформации полипептидной цепи.

[2] Водородные связи между пептидами, разделенными тремя парами углов ϕ и ψ, играют важную роль в стабилизации α-спиральных конформаций вблизи минимумов областей I и II. Этот факт не был учтен при построении рис. 5.7 и 5.8. Следовательно, эти рисунки отражают ситуации, в которых влияние растворителя или какие-нибудь другие факторы исключают возможность образования водородных связей.

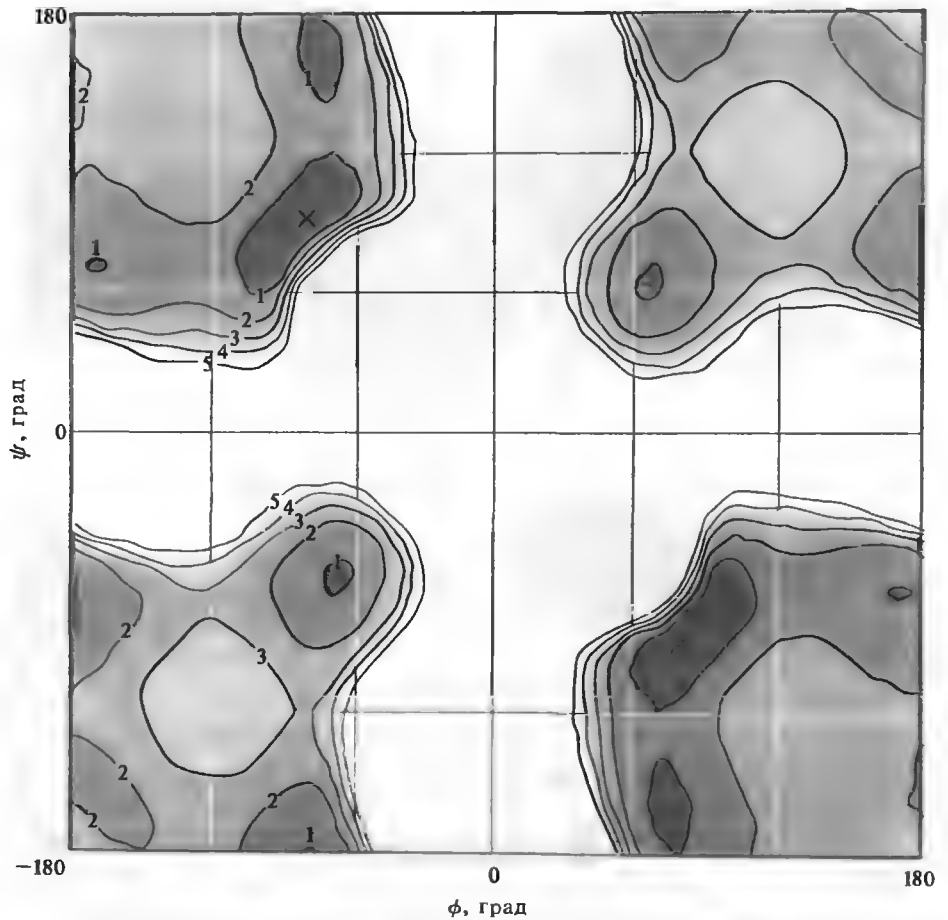


РИС. 5.7. Энергетическая контуриая карта для остатка глицина в полипептидной цепи, рассчитанная согласио уравнению (5.5). Градации цвета соответствуют 1 ккал. Крестиками обозначены самые низкие значения энергии. [D.A. Brant, P.J. Flory, J.Mol.Biol., **23**, 47 (1967).]

цепям полиаланина. Эти измерения показали, что тенденцию скелета цепи находиться в вытянутых конформациях можно объяснить только учетом электростатических взаимодействий в нем. (Этот вопрос обсуждается более детально в гл. 18.)

ВЫВОДЫ ИЗ АНАЛИЗА КОНФОРМАЦИОННОЙ ЭНЕРГИИ ГЛИЦИЛА И L-АЛАНИЛА

Анализ энергетических карт глицина и аланина может привести к весьма интересным выводам. Из-за наличия симметрично расположенных низкоэнергетических областей глициновые остатки, естественно, обладают большей гибкостью, чем аланин или подобные ему (содержащие боковые цепи) аминокислоты. Более того, области, которые предпочитает глицин, расположенный ближе к центру энергетической карты ($\phi = 0°$, $\psi = 0°$), чем в

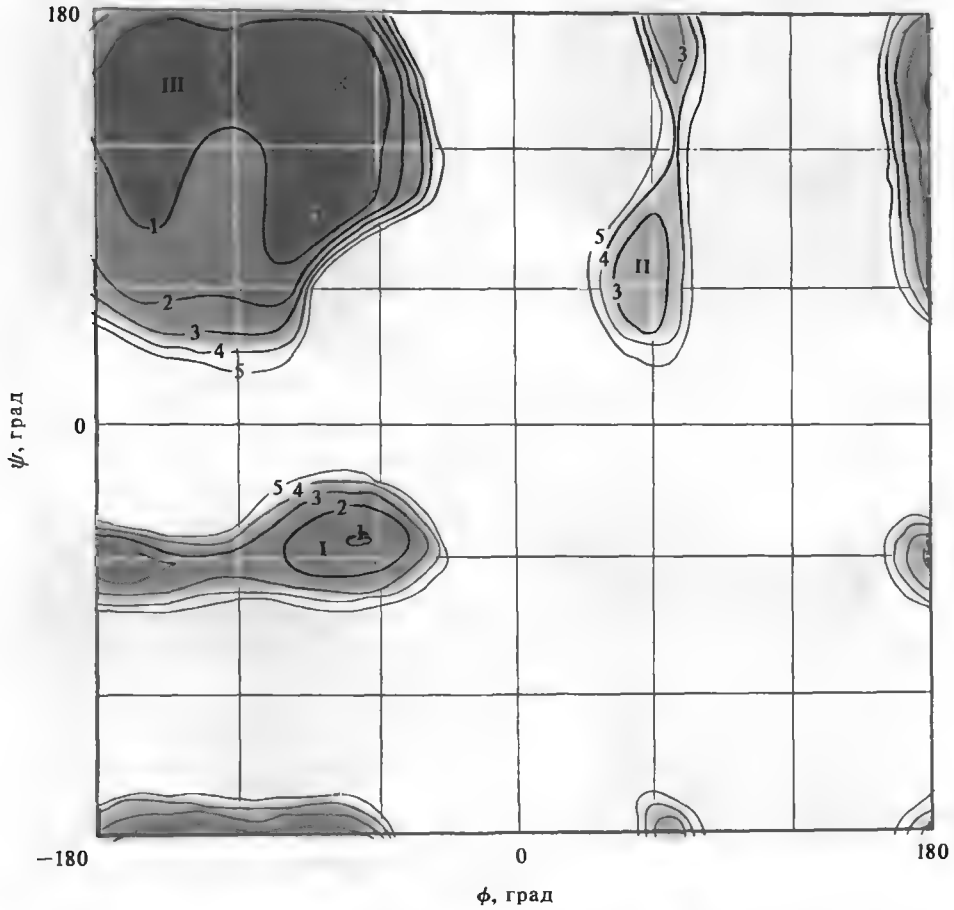


РИС. 5.8. Энергетическая контурная карта для остатка L-аланина в полипептидной цепи, рассчитанная согласно уравнению (5.5). [D.A. Brant, P.J. Flory, J.Mol.Biol., **23**, 47 (1967).]

случае аланина. Благодаря этому глицин содействует сворачиванию полипептидной цепи в более компактные структуры. Это предсказание подтверждается тем, что среднеквадратичное расстояние между концами полиглицинового статистического клубка, состоящего из 100 остатков, почти в два раза меньше, чем в случае полиаланинового статистического клубка из 100 остатков. Кроме того, введение нескольких остатков глицина в полиаланиновую цепь вызывает непропорционально большое уменьшение ее размеров (см. конформационную статистику в гл. 18).

КОНФОРМАЦИОННАЯ ЭНЕРГИЯ НЕКОТОРЫХ ДРУГИХ ОСТАТКОВ

L-аминокислотные остатки, не разветвляющиеся у β-углеродного атома, имеют энергетические карты, весьма похожие на карту L-аланилового остатка (об этом упоминалось выше в связи со стерическими картами). Сходное расположение трех низкоэнергетических областей имеют и остатки типа валина с разветвляющейся у β-углеродного атома боко-

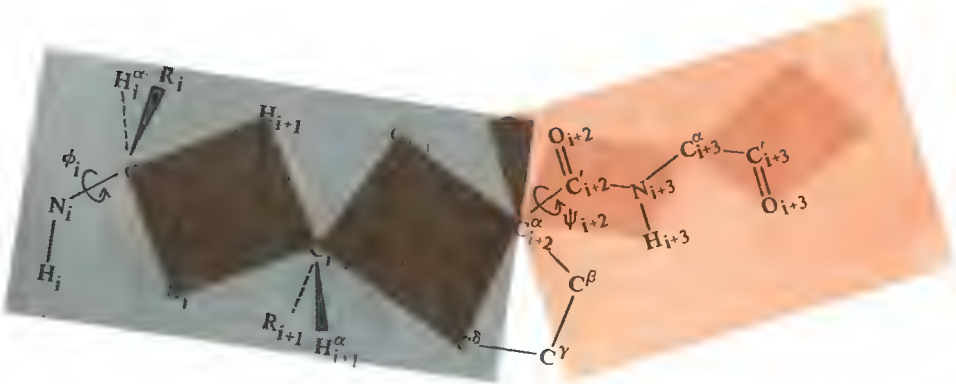


РИС. 5.9. Схематическое изображение полипептидной цепи, содержащей изолированный остаток пролина. [P.R. Schimmel, P.J. Flory, J.Mol.Biol., **34**, 105 (1968).]

вой цепью, но их площади несколько меньше. Однако один остаток, L-пролил, совершенно отличен от других и должен рассматриваться отдельно. Кроме того, остаток, который предшествует пролину в полипептидной цепи, также заслуживает особого обсуждения.

ИЗОЛИРОВАННЫЙ ОСТАТОК L-ПРОЛИЛ

В большинстве белков остаток пролина встречается в цепи изолированно, т.е. последовательности из двух и большего числа пролинов отсутствуют или крайне редки. Вследствие того что пролин с его объемным пирролидиновым кольцом ограничивает область возможных конформаций предшествующего остатка, конформационные свободы изолированного пролина и пролина, внедренного в последовательность таких же остатков, значительно различаются. Для нас изолированный пролин более интересен, и мы ограничимся рассмотрением этого случая.

На рис. 5.9 представлена схема α-L-полипептидной цепи, содержащей изолированный остаток пролина в *транс*-конфигурации[1]. Жесткое пирролидиновое кольцо фиксирует угол поворота на значении, приблизительно равном $-60°$. Поэтому конформационная энергия пролила зависит только от угла ψ.

На рис. 5.10 представлена зависимость конформационной энергии изолированного остатка пролина от угла поворота ψ. Оба минимума при $\psi \approx -55°$ и $\psi \approx 145°$ попадают в области I и III рис. 5.8 при $\phi = -60°$. Амид-имидное дипольное взаимодействие в расчеты не включалось вследствие трудности его оценки; самое большее, что оно может, — это понизить относительное положение минимума при $\psi \approx 145°$ на 1 ккал/моль.

Из двух возможных конформаций, разрешенных для изолированного остатка пролина, одна соответствует относительно компактной форме, другая — относительно растя-

[1] Известно, что пролин существует в полипептидах и в *цис*-форме, однако данные рентгеноструктурного анализа показывают, что в белках *транс*-форма встречается значительно чаще.

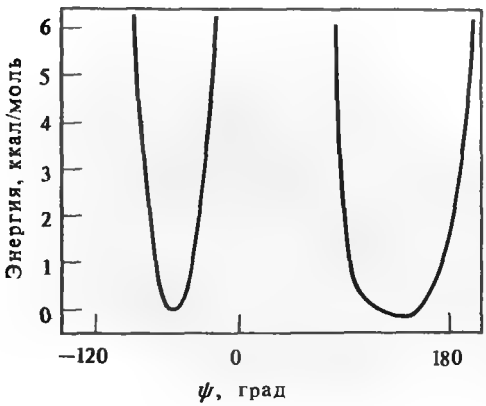


РИС. 5.10. Зависимость конформационной энергии от угла ψ для изолированного остатка пролина в полипептидной цепи. [P.R. Schimmel, P.J. Flory, J.Mol. Biol., **34**, 105 (1968).]

нутой. Разность энергий этих двух состояний очень мала, так что остаток пролина может находиться в любом из них. Относительная легкость, с которой остаток пролина принимает компактную конформацию ($\psi \approx -55°$), делает его желательным кандидатом для создания изгибов и поворотов в полипептидной цепи.

КОНФОРМАЦИОННАЯ ЭНЕРГИЯ ОСТАТКА, ПРЕДШЕСТВУЮЩЕГО L-ПРОЛИЛУ

Мы уже упоминали, что повороты на углы ϕ_i и ψ_i i-го остатка в полипептидной цепи не зависят от поворотов в соседних $(i-1)$ и $(i+1)$-остатках. Следовательно, стерические взаимодействия остатка i с предшествующим остатком $i-1$ не зависят от природы этого остатка, так как в них участвует только его карбонильная группа (C = O). Однако взаимодействия i-го и $(i+1)$-го остатков зависят от того, не является ли последний остатком пролина. Если последующий остаток — пролил, обычные взаимодействия атомов i-го остатка с амидным водородом $i+1$-го заменяются взаимодействиями с атомами пирролидинового кольца, в частности с $-CH_2$-группой, связанной с азотом имида, а на эти взаимодействия накладываются стерические ограничения. Поэтому можно ожидать, что вращение остатка, за которым следует пролин, будет менее свободным, чем в тех случаях, когда за ним следует любой другой остаток.

На рис. 5.11 представлена энергетическая карта для остатка L-аланина, предшествующего остатку L-пролина. Сравнение рис. 5.11 и 5.8 показывает, что присутствие пролина приводит к значительному сужению области разрешенных для аланина конформаций. Область I становится запрещенной, так как при повороте на угол $-180° \leqslant \psi \leqslant 60°$ возникают стерические препятствия. Эти дополнительные энергетически невыгодные взаимодействия в основном объясняются наложением боковой группы аланина $-CH_3$ и группы пролина $-CH_2$, связанной с азотом имидной группы.

В случае глицина, предшествующего L-пролину, энергетическая карта мало отличается от представленной на рис. 5.7. Все четыре разрешенные области остаются неизменными, но энергия в окрестности $\psi = 0°$ увеличивается вследствие появления энергетически невыгодных взаимодействий, в которых принимают участие атомы пирролидинового кольца.

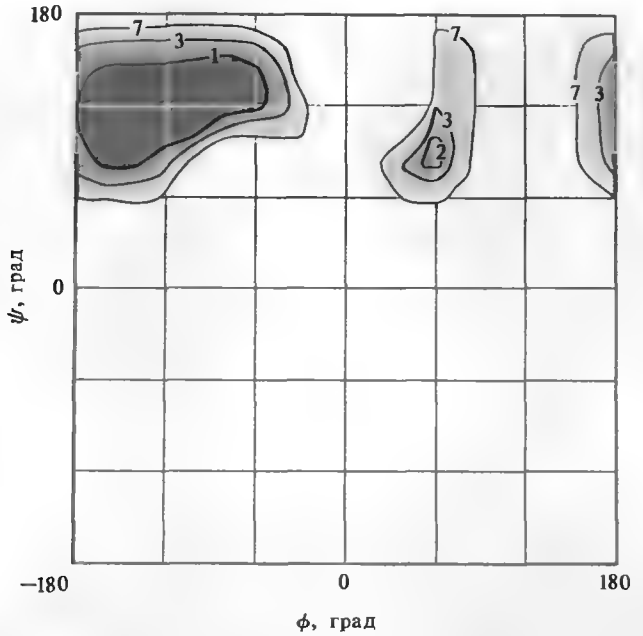


РИС. 5.11. Энергетическая контурная карта для остатка L-аланина, за которым следует остаток пролина в полипептидной цепи. [P.R. Schimmel, P.J. Flory, J.Mol.Biol., **34**, 105 (1968).]

КОНФОРМАЦИОННЫЕ ОГРАНИЧЕНИЯ, НАЛАГАЕМЫЕ НА ОСТАТОК, ПРЕДШЕСТВУЮЩИЙ ПРОЛИЛУ

Область I запрещена для остатков с боковой цепью, предшествующих пролину. Этот факт имеет очень важное следствие: такие остатки не могут участвовать в формировании правой α-спирали. С другой стороны, для самого остатка пролина это ограничение отсутствует (см. рис. 5.10), даже если его азот и не способен участвовать в образовании водородной связи. Поэтому пролин может встречаться только в начале α-спиральной последовательности, но «не позволяет» предшествующему остатку, если это не глицил, находиться в составе α-спирали. Таким образом, то обстоятельство, что азот пролина не может участвовать в образовании водородной связи, не является основной причиной способности этого остатка дестабилизировать α-спираль, как часто считают.

5.5. Экспериментально наблюдаемые значения углов поворота

Выше мы рассмотрели энергетически разрешенные конформации аминокислотных остатков, расположенных внутри полипептидных цепей. Для построения энергетических карт можно было воспользоваться, разумеется, только полуэмпирическими энергетическими функциями. Естественно, возникает вопрос, насколько общие предсказания, сделанные на основе этих расчетов, согласуются с наблюдаемыми конформациями амино-

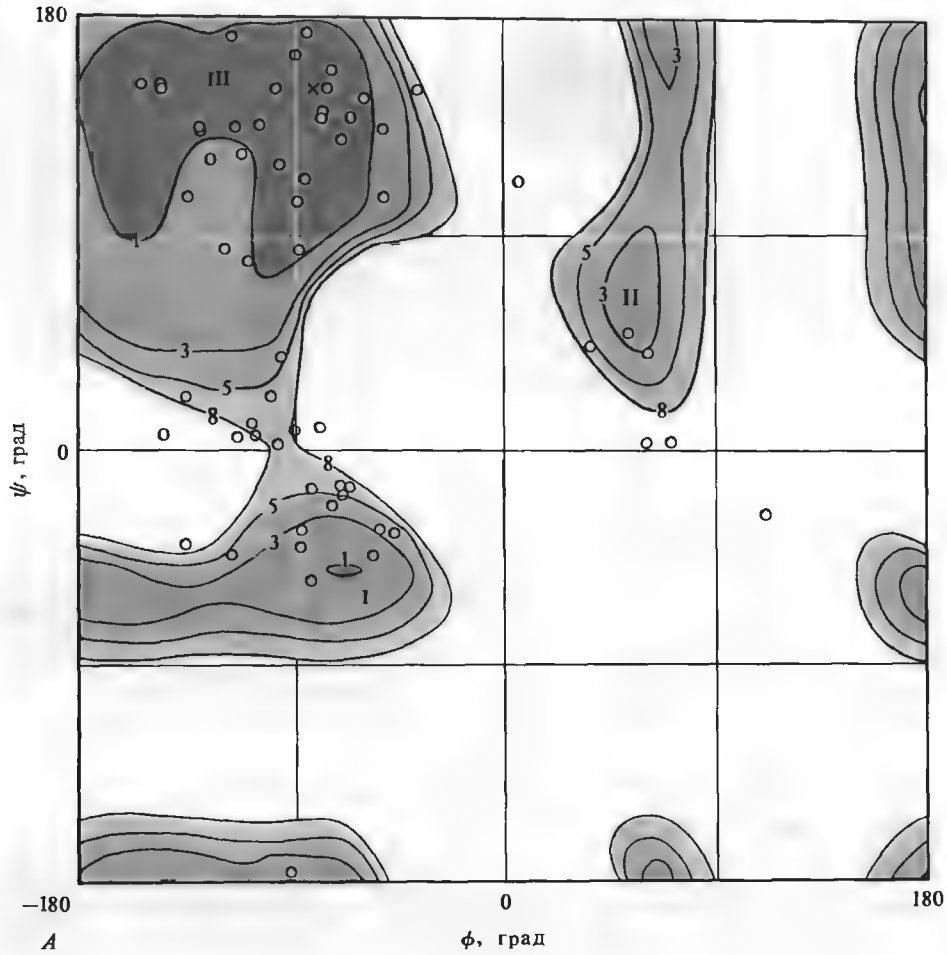
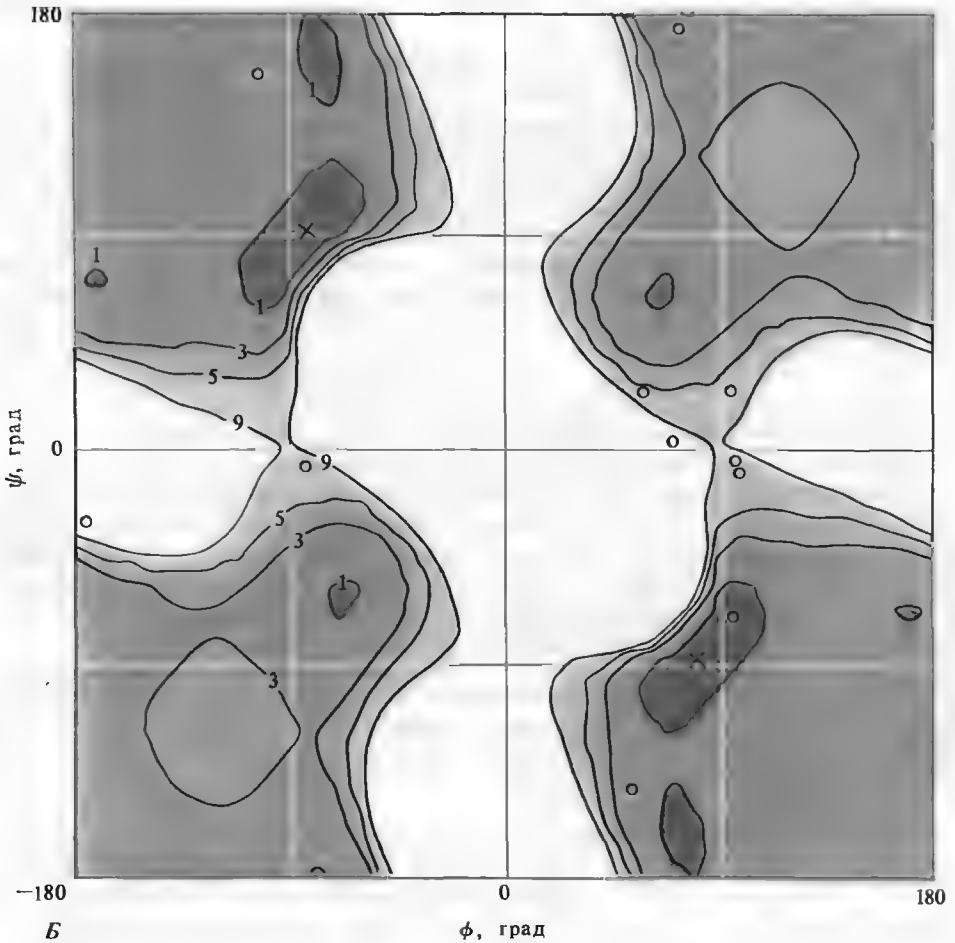


РИС. 5.12. Наблюдаемые координаты остатков, нанесенные на конформационные энергетические карты. Кружками отмечены координаты остатков, обнаруженных в лизоциме белка куриного яйца. Конформационные энергетические карты рассчитывались для остатков L-аланина (*А*) и глицина (*Б*) внутри полипептидной цепи. [D.A. Brant, P.R. Schimmel, Proc. Natl. Acad. Sci. USA, **58**, 428 (1967).]

кислотных остатков. Наилучший ответ может дать анализ углов ϕ и ψ в различных кристаллических структурах пептидов, полипептидов и белков.

Нужно, однако, предостеречь читателя. Есть несколько факторов, которые могут вызывать или увеличивать расхождение между предсказанными и наблюдаемыми конформациями. Один из них вытекает из предположения о фиксированных длинах связей и величинах валентных углов, лежащего в основе построения карт, приведенных на рис. 5.3, 5.4, 5.7, 5.8, 5.10 и 5.11. В то же время вполне возможны как деформация валентных углов на несколько градусов, так и небольшие изменения длин связей; и то, и другое может вызвать изменение формы «разрешенных» областей. Кроме того, ошибка, с которой определяются углы ϕ и ψ в кристаллах белков, составляет по крайней мере несколько градусов.



Б

ПРИМЕР ЛИЗОЦИМА

Учитывая все эти соображения, мы можем продолжить обсуждение вопроса о том, насколько хорошо наблюдаемые конформации согласуются с энергетическими картами. Рассмотрим, например, остатки кристаллического лизоцима, которые расположены в неспиральных областях этой белковой молекулы. (Остатки, входящие в состав спиральных участков, принимают участие во взаимодействиях, которые мы не учитывали при расчетах конформационной энергии.[1])

На рис. 5.12, *А* представлена энергетическая контурная карта для остатка, содержащего боковую цепь CH_2-R' у α-углеродного атома и расположеиного внутри полипептид-

[1] Например, в α-спиралях между группами, разделенными тремя парами углов ϕ и ψ, образуются водородные связи. Эти связи ие обязательно сохраняются в водном окружении, где молекулы воды могут конкурировать с центрами, образующими водородные связи в белке. В свернутом белке, конечио, цепь сама может формировать «карманы», создающие неполярное окружение, благоприятиое для образования внутримолекулярных водородных связей.

цепи. Кружками отмечены положения отдельных остатков. (Эта карта идентична карте для аланина, приведенной на рис. 5.8.)

Легко видеть, что почти все остатки попадают в области, в которых энергия не особенно велика. В самом деле, 22 остатка (из 61, отмеченного на рис. 5.12, *А*) расположены в окружающей минимум зоне, в которой энергия не превышает 1 ккал/моль. Вообще говоря, согласно принципам статистики, относительное количество остатков, попадающих в ту или иную область, зависит как от энергии, так и от эффективной площади этой области. Исходя из подобных качественных соображений, можно ожидать, что порядок относительной заселенности трех низкоэнергетических зон будет выглядеть следующим образом: III > I > II. Именно это и наблюдается. Однако видно также, что вблизи седловины между областями I и III, где энергия относительно велика, расположено больше остатков, чем в области II. Такое кажущееся противоречие можно объяснить действием факторов, рассмотренных выше.

Степень несовпадения структуры некоторых остатков с рассчитанными преимущественными конформациями может служить мерой величин тех сил, которые не были учтены при расчетах. Стерические или другие взаимодействия между остатками, расположенными далеко друг от друга в аминокислотной последовательности, взаимодействия белка с растворителем и, наконец, силы, возникающие при кристаллической упаковке молекул, по-видимому, могут быть достаточно велики для того, чтобы побудить остаток принять конформацию, находящуюся на границе или даже за пределами «разрешенной» области. Поэтому некоторые расхождения между рассчитанными и наблюдаемыми структурами могут быть связаны не с несовершенством расчетов или данных рентгеноструктурного анализа, а с тем, что не были учтены эти силы.

На рис. 5.*12,Б* представлена энергетическая контурная карта для остатка глицина (ср. с рис. 5.7), расположенного внутри полипептидной цепи; кружками отмечены 11 конформаций этого остатка, в которых он находится в лизоциме. Две конформации попадают в области с чрезвычайно высокой энергией. Остальные рассеяны по «разрешенным» областям и не имеют тенденции группироваться в наиболее низкоэнергетической (1 ккал/моль) зоне.

Подобное сравнение предсказанных и полученных экспериментально координат ϕ и ψ отдельных остатков было проведено для других белков и различных полипептидных структур. По большей части эти остатки попадают в «разрешенные» области энергетических карт. Таким образом, пытаясь предсказать конформацию того или иного белка, мы можем с достаточным основанием полагать, что конформации отдельных остатков находятся в пределах «разрешенных» зон. Однако диапазон возможных конформаций в самих этих зонах достаточно велик, чтобы допустить огромное число конформаций полимера. Поэтому следует обратить внимание на специфические силы, которые помогают стабилизировать высокоупорядоченные структуры нативных белковых молекул. Именно эти силы из всех стерически разрешенных конформаций выделяют одну (или несколько), наиболее выгодную в термодинамическом отношении.

Из числа специфических сил и взаимодействий, определяющих нативную трехмерную структуру белков, можно выделить несколько, роль которых особенно велика. Это ковалентные связи между сегментами цепи, осуществляемые с помощью дисульфидных мостиков, нековалентные взаимодействия, такие, как водородные связи и ионные силы. Наконец, к ним относятся «гидрофобные» взаимодействия и другие взаимодействия макромолекулы с растворителем. Все они вместе взятые достаточно сильны и специфичны для того, чтобы заставить молекулу принять наиболее предпочтительную конформацию.

5.6. Водородные связи

Обычно говорят, что две молекулы, донор D—H и акцептор A, образуют водородную связь, когда имеются свидетельства того, что в образовании ассоциата между ними особую роль играет атом водорода донора. Выявить наличие водородных связей в тех системах, где их образование в принципе возможно, помогают спектроскопические методы. Например, этот процесс сопровождается сдвигом инфракрасных полос и полос спектров комбинационного рассеяния. Еще один прямой физический метод, которым можно воспользоваться для обнаружения водородных связей, — протонный магнитный резонанс (см. гл. 9). Вместе с тем подтверждение существования тесных межмолекулярных ассоциатов можно получить путем сравнения обычных макроскопических свойств — таких, как температура замерзания и кипения чистых жидкостей, а также теплоты парообразования. В табл. 5.4 приведены температуры плавления и кипения для нескольких веществ, молекулы которых имеют близкие размеры.

Среди прочего отметим, например, относительно высокие по сравнению с H_2S температуры замерзания и кипения воды. Огромное различие в этих макроскопических свойствах ясно указывает на сильные межмолекулярные взаимодействия в воде. Рассмотрим также температуры кипения трех весьма схожих ароматических жидкостей: хлорбензол кипит при 132°C, о-хлорфенол — при 175°C и *п*-хлорфенол — при 217°C. Как объяснить эти различия?

Таблица 5.4

СРАВНЕНИЕ ТЕМПЕРАТУР ПЛАВЛЕНИЯ И КИПЕНИЯ ВЕЩЕСТВ, МОЛЕКУЛЫ КОТОРЫХ БЛИЗКИ ПО РАЗМЕРУ[1]

Соединение	Температура плавления, К	Температура кипения, К	Соединение	Температура плавления, К	Температура кипения, К
H_2O	273	373	CH_3CH_2OH	156	351
H_2S	190	211	$CH_3CH_2CH_3$	83	231
$CH_3-\overset{\overset{O}{\|}}{C}-OH$	290	391	CH_3-O-CH_3	135	249
$CH_3-\overset{\overset{O}{\|}}{\underset{H}{C}}-CH_3$	178	330	CH_3NH_2	181	267
			CH_3CH_3	101	185

[1] G. C. Pimentel, A. L. McClellan, The Hydrogen Bond (San Francisco: W. H. Freeman and Company, 1960), p. 36; Handbook of Chemistry and Physics, 39th ed. (Cleveland, Ohio: Chemical Rubber Pub. Co., 1958).

Экспериментально обнаружено, что водородная связь между D—H и A может формироваться, когда D и A являются сильно электроотрицательными. В большинстве случаев все три атома при образовании связи должны быть коллинеарны. Каковы же физические принципы образования водородной связи? Раньше считали, что эта связь имеет в осиовном электростатическую природу и образуется в результате кулоновского притяжения. Полагали, что небольшое разделение парциальных зарядов на атомах, приводящее к кулоновскому взаимодействию, существует и до образования связи. Эта точка зрения подтверждалась тем, что атомы фтора образуют наиболее прочные водородные связи, атомы кислорода — несколько менее прочные, атомы азота — слабые. Однако такой простой электростатический подход не может объяснить всех экспериментальных фактов, например, отсутствие всякой связи между величиной дипольного момента и прочностью

водородной связи в основаниях. Ясно, что здесь требуется более сложное квантовомеханическое рассмотрение, и было предложено несколько его вариантов. Но электростатическая модель все еще остается полезной для получения разного рода качественных оценок.

КОНКУРЕНЦИЯ МОЛЕКУЛ ВОДЫ ЗА МЕСТА ОБРАЗОВАНИЯ ВОДОРОДНЫХ СВЯЗЕЙ В БЕЛКЕ

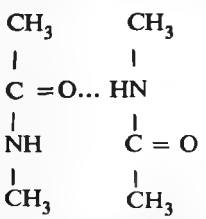
Для формирования белковой структуры наиболее важны связи между атомами азота амидных групп и атомами кислорода карбонильных групп скелета. Эти связи играют основную роль в стабилизации различных спиральных и других упорядоченных полипептидных конформаций (таких, как α-спираль и β-слой); огромна их роль также в стабилизации упорядоченных участков свернутых белковых молекул. Поэтому многие исследователи предположили, что водородные связи должны играть решающую роль в организации конформации белков.

Выше уже указывалось, что вода — растворитель нативных белков — охотно участвует в формировании водородных связей. Следовательно, уместен вопрос об изменении свободной энергии в реакции

$$NH\cdots H_2O + OH_2\cdots O{=}C \rightleftarrows NH\cdots O{=}C + \text{свободная } H_2O \qquad (5.6)$$

Ответ на этот вопрос совсем не очевиден; особенно его осложняет то обстоятельство, что освобождающиеся молекулы воды, вероятно, образуют новые водородные связи с другими молекулами воды.

Проблема в целом подробно исследовалась Клотцом и Франценом (1962 г.) методом инфракрасной спектроскопии. Они изучали ассоциацию молекулы N-метилацетамида (NMA) — достаточно близкого аналога пептидного скелета. Молекулы NMA могут формировать водородную связь с участием амидной группы следующим образом:



причем амидная группа находится в *транс*-конформации. Изучалось равновесие между мономерами (M) и димерами (D) и проводился анализ степени ассоциации α, определяющейся следующим образом:

$$\alpha = 2(D)/[(M) + 2(D)] = \text{доля молекул в димерном состоянии}$$

Параметр α измерялся в трех различных растворителях — четыреххлористом углероде, диоксане и воде, что позволяло прямо проверить влияние растворителя. На рис. 5.13 показана зависимость α от концентрации NMA, полученная в результате этих экспериментов. Совершенно очевидно, что степень ассоциации молекул NMA в трех разных растворителях заметно различается. Наиболее сильная ассоциация происходит в четыреххлористом углероде, который сам не образует водородных связей. Значительно меньше степень ассоциации в воде — сильном конкуренте за места образования водородных связей в NMA. Степень ассоциации NMA в диоксане, молекулы которого имеют два атома кислорода, способных образовывать водородную связь, имеет промежуточное значение.

Для оценки термодинамических параметров была также измерена температурная зависимость степени ассоциации во всех трех растворителях (табл. 5.5). Очевидно, что в че-

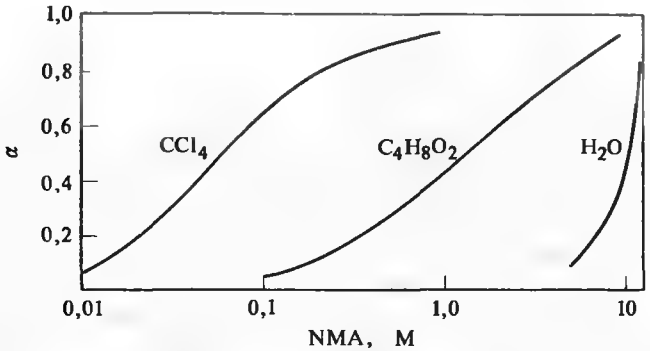


РИС.5.13. Ассоциация NMA в четыреххлористом углероде, диоксане и воде. [I.M. Klotz, J.S. Franzen, J. Am. Chem. Soc., **84**, 3461 (1962).]

тыреххлористом углероде ассоциация молекул NMA термодинамически более выгодна, чем в воде.

Эти результаты явно указывают на необходимость учета взаимодействий с растворителем при определении наиболее выгодного состояния белковых молекул. Как этот фактор влияет на наши представления о прочности водородных связей в белках? Внутренняя часть белковой молекулы в основном имеет углеводородную природу, и для поиимания

Таблица 5.5

ТЕРМОДИНАМИКА ОБРАЗОВАНИЯ ВОДОРОДНОЙ СВЯЗИ МЕЖДУ МОЛЕКУЛАМИ N-МЕТИЛАЦЕТАМИДА С УЧАСТИЕМ АМИДНОЙ ГРУППЫ ПРИ 25°C[1)]

Растворитель	ΔH^0, ккал · моль$^{-1}$	ΔS^0, кал · град$^{-1}$ · моль$^{-1}$	ΔG^0, ккал · моль$^{-1}$
CCl_4	−4,2	−11	−0,92
Диоксан	−0,8	−4	0,39
H_2O	0,0	−10	3,1

[1)] I. M. Klotz, J. S. Franzen, J. Am. Chem. Soc., **84**, 3461 (1962). Стандартиая концентрация 1 М.

процессов образования и разрыва водородных связей в этой области, по-видимому, наиболее пригодны данные, полученные в четыреххлористом углероде. Вместе с тем результаты измерений, проведенных в воде, должны лучше соответствовать случаю, когда водородные связи образуются вблизи полярной поверхности молекулы белка или в местах, доступных для воды. Нельзя, однако, упускать из виду, что наличие у воды особой структуры вблизи поверхности белка еще более усложняет ситуацию. Именно на вопросы, связанные со структурой воды и ее ролью в формировании конформаций белков, мы обратим сейчас свое внимание.

5.7. Гидрофобные взаимодействия и структура воды

ФИЗИЧЕСКИЕ СВОЙСТВА ВОДЫ

Уже давно известно, что вода — один из основных компонентов биологических систем — обладает весьма необычными свойствами. В действительности вода представляет собой столь сложный объект, что ее изучение превратилось в одну из наиболее трудных и интересных областей современной экспериментальной и теоретической физической химии. Примером особых физических свойств воды может служить хорошо известное увеличение ее объема при замерзании — явление, прямо противоположное тому, что наблюдается у большинства жидкостей. Это увеличение удельного объема составляет 1,6 $см^3$/моль. Более того, когда лед при постепенном нагревании плавится при температуре 0° С, образовавшаяся жидкость продолжает сжиматься даже после того, как он весь растает. Это уменьшение объема происходит до температуры 4° С, выше которой удельный объем воды начинает увеличиваться. Следовательно, максимальная плотность воды достигается при температуре 4° С. Помимо этих особых макроскопических свойств вода обладает аномально высокой температурой плавления и кипения, а также необычно высокой теплоемкостью.

Мощным методом исследования структуры жидкостей служит инфракрасная спектроскопия. Однако инфракрасный спектр воды невероятно сложен, и большинство полос поглощения пока не удается объяснить количественно. Поэтому данные такого рода в настоящее время не могут дать детальной информации о структуре воды.

На рис. 5.14 схематически показана структура кристаллической воды. Каждая молекула участвует в образовании четырех водородных связей, так что образуется тетраэдрическая структура. В двух из этих связей атом кислорода является акцептором, две другие об-

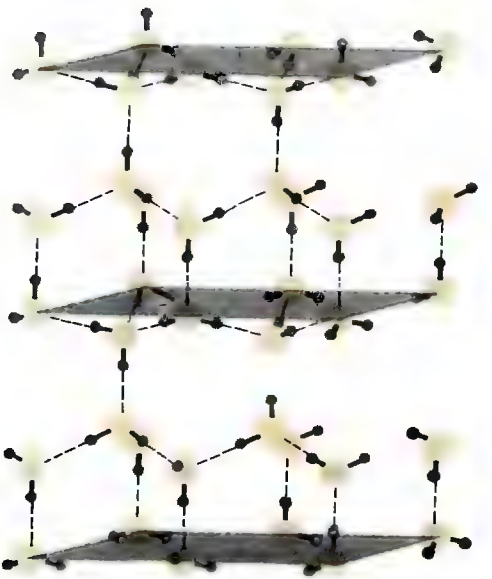


РИС. 5.14. Схематическая модель кристаллической воды. Пунктиром показаны водородные связи между молекулами воды. [L. Pauling, The Nature of the Chemical Bond, 3-d ed. (Ithaca, N.Y.: Cornell Univ. Press, 1960).]

разуются между атомами водорода и соседними молекулами. Могут ли водородные связи *в жидкости* не играть ключевой роли? Определяющая роль этих связей в структуре воды признается повсеместно, хотя конкретное объяснение того, как водородные связи влияют на микроструктуру, может меняться от одной теории к другой.

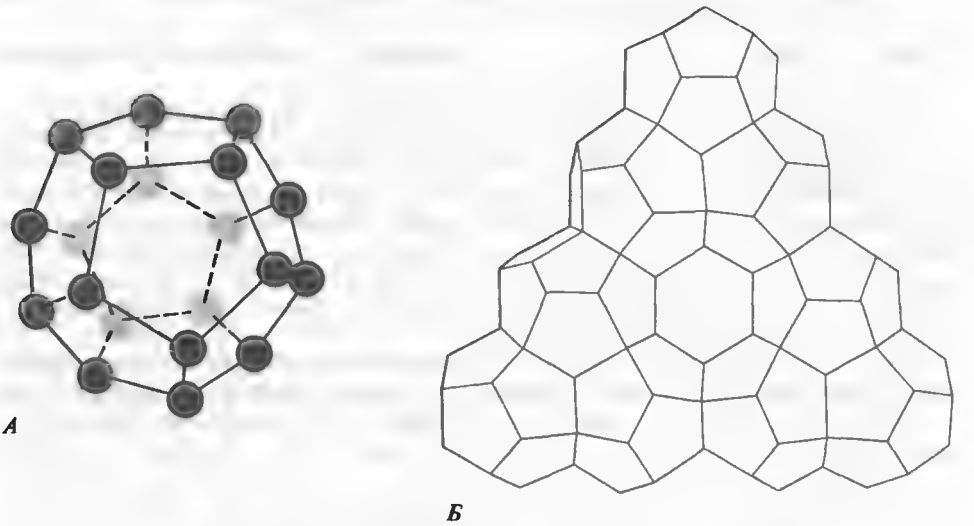
Весьма распространено качественное представление, согласно которому молекулы воды в растворе образуют различного размера кластеры с помощью водородиых связей, причем эти кластеры постоянно распадаются и вновь образуются (см. Frank, Wen, 1957). Они напоминают кристаллические айсберги, хотя картина водородных связей и геометрия упаковки могут отличаться от таковых для кристаллов льда; например, молекулы, возможно, образуют меньше четырех связей. Между кластерами разной величины расположены несвязанные молекулы воды, которые обладают большей свободой перемещения и более плотной упаковкой, так как никакие силы не вынуждают их сохранить дистанции, характерные для кристаллических решеток.

Такая качественная картина помогает объяснить некоторые макроскопические свойства воды. Например, уменьшение объема при таянии льда может происходить из-за освобождения части ее молекул из решетки, скрепленной водородными связями. Эти «свободные» молекулы способны упаковываться более компактно, что и приводит к уменьшению объема при переходе из твердого состояния в жидкое. По мере возрастания температуры от 0° С все большее число молекул освобождается и вода продолжает сжиматься. Однако с ростом температуры увеличиваются тепловые колебания молекул, что приводит к увеличению объема. В результате влияния этих двух противоположных факторов максимальная плотность достигается при температуре 4° С, выше которой вода при нагревании только расширяется.

ВЛИЯНИЕ РАСТВОРИТЕЛЯ НА СТРУКТУРУ БЕЛКА

Как же связаны изложеннные соображения со структурой нативного белка? Иметь представление о структуре растворителя и о природе его взаимодействия с белком необходимо для того, чтобы понять, какое огромное влияние может оказывать растворитель на конформацию белковых молекул. То обстоятельство, что относительно низких концентраций мочевины или гуанидинхлорида (например, 1 молекула мочевины на 5 ÷ 10 молекул H_2O) часто оказывается достаточно для полного разрушения нативной структуры белка и перехода его в неупорядоченную конформацию, хорошо иллюстрирует чувствительность структуры белковых молекул к составу раствора. С другой стороны, растворитель может вызывать структурные изменения, приводящие к формированию иного, уникального, высокоупорядоченного состояния. Для этой цели, например, может быть использован 2-хлорэтанол. Белки растворимы в нем, и хотя это полярный растворитель, он обладает гораздо более слабой, чем вода, способностью образовывать водородные связи. Например, рибонуклеаза в 2-хлорэтаноле принимает другую, как полагают, значительно более совершенную спиральную конформацию, чем в воде (см. Doty, 1960). Сходные результаты были получены и для других белков. Их можно объяснить тем, что в «бедных» водородными связями растворителях легче образуются внутримолекулярные водородные связи. Все это еще раз подчеркивает, как велико влияние растворителя на конформацию.

Ясно, что вода может влиять на конформацию белка, конкурируя за места образования водородных связей и эффективно сольватируя полярные группы. Но этими взаимодействиями, однако, не исчерпывается все влияние воды на конформацию. Например, мочевина также может участвовать в образовании водородных связей и взаимодействовать с полярными группами, но, когда она присутствует в достаточных количествах, она действует как денатурирующий агент. Для лучшего понимания роли воды рассмотрим, как водное окружение взаимодействует с неполярными боковыми группами.



А *Б*

РИС. 5.15. Додекаэдры. *А*. Додекаэдр, образованный молекулами воды (кружки) и содержащий полость диаметром около 5 Å. [M.V. Stackelberg, H.R. Müller, Z. Electrochem., **58**, 25 (1954).] *Б*. Кластер додекаэдров. [A.F.Wells, The Third Dimension in Chemistry (Oxford: Oxford Univ. Press. 1956).]

НЕПОЛЯРНЫЕ МОЛЕКУЛЫ В ВОДЕ

Давно известно, что неполярные молекулы образуют в воде очень стабильные кристаллические гидраты, в которых похожие на клетки многогранники, состоящие из молекул воды, формируют полости, окружающие неполярные молекулы. Эти многогранники могут объединяться, образуя пространственную решетку. На рис. 5.15, *А* показана структура пентагонального додекаэдра, в котором отдельные кружочки представляют молекулы воды. Полость внутри додекаэдра имеет диаметр приблизительно 5 Å — достаточно

Таблица 5.6

ТЕПЛОТА ГИДРАТАЦИИ НЕПОЛЯРНЫХ СОЕДИНЕНИЙ[1)]

Соединение	$-\Delta H$, ккал/моль	Соединение	$-\Delta H$, ккал/моль
Ar	16,6	C_2H_4	15
Kr	13,9 – 16,5	C_2H_6	15
CH_4	14,5 – 17	CH_3SH	16,6
C_2H_2	15		

[1)] I. M. Klotz, Brookhaven Symp. Biol., **13**, 25 (1960).

большой, чтобы вместить соответствующую неполярную молекулу. На рис. 5.15, *Б* показан кластер додекаэдров.

Такие гидраты обладают замечательной стабильностью — так, при растворении в воде неполярных молекул до достаточно высоких концентраций температура замерзания смеси поднимается значительно выше 0° С. Гидраты характеризуются также большой по абсолютной величине отрицательной теплотой образования. В табл. 5.6 приведены значения теплоты образования нескольких иеполярных гидратов. Несмотря на большое различие между молекулами в размерах и составе, все теплоты гидратации близки по величине — около -15 ккал/моль. Отсюда следует, что большой выигрыш в энтальпии обуславливается не непосредственно взаимодействиями неполярных молекул с растворителем, а освобождением энергии при организации молекул воды в близкие по своим размерам, похожие на клетки структуры.

ВКЛАДЫ АДДИТИВНОЙ ЭНТРОПИИ И ЭНТРОПИИ СМЕШЕНИЯ В СВОБОДНУЮ ЭНЕРГИЮ ПЕРЕНОСА

Подход к рассмотрению взаимодействий неполярных молекул с водой был разработан В. Козманом. Не фиксируя свое внимание на образовании кристаллических гидратов, он рассмотрел термодинамику переноса неполярных остатков из неполярного же растворителя в воду. Например, он задался вопросом, какова свободная энергия переноса мольной доли χ метана из бензола в воду. Цель такого подхода — определить, предпочитают ли неполярные группы располагаться на поверхности белка, чтобы взаимодействовать с водой, или они стремятся собраться вместе внутри белковой молекулы, чтобы создать для себя неполярное окружение.

Прежде чем рассмотреть данные по свободной энергии переноса, необходимо обсудить различие между вкладами *аддитивной энтропии* и *энтропии смешения* в свободную энергию (см. Gurney, 1953; Kauzmann, 1959). Мы должны быть уверены, что все результаты рассматриваются на основе одинаковых общих соображений. Чтобы проиллюстрировать встречающиеся трудности, обратимся к реакции

$$A + B \rightleftarrows AB \tag{5.7}$$

характеризующейся константой равновесия

$$K = (AB)/(A)(B) \tag{5.8}$$

и изменением стандартной свободной энергии

$$\Delta G^0 = -RT \ln K \tag{5.9}$$

Если в качестве единицы концентрации мы выберем моль/л, то ΔG^0 будет равно свободной энергии в том случае, когда стандартные состояния веществ A, B и AB соответствуют концентрациям 1 моль/л. Однако численное значение K (и, следовательно, ΔG^0) будет иным при других стандартных состояниях (или при другом выборе единицы измерения концентрации). Конечно, свободная эиергия зависит от выбора стандартного состояния только тогда, когда число реагентов не равно числу продуктов реакции.

Эта зависимость представляет собой чисто статистический эффект, обусловленный вкладом парциальной молярной энтропии. Рассмотрим S_A — парциальную молярную энтропию вещества A:

$$S_A = S'_A - R \ln \chi_A \tag{5.10}$$

где χ_A — мольная доля вещества A (см. Дополнение 5.2). (Мы рассматриваем разбавленные растворы и поэтому пренебрегаем влиянием коэффициента активности.) Величина $-R \ln \chi_A$ называется *энтропией смешения*. Этот чисто статистический член характеризу-

ет эффект смешивания вещества А с растворителем и не зависит от его структуры и свойств. Химическую природу вещества А и его взаимодействие с растворителем учитывает другой член выражения (5.10), S'_A, называемый *аддитивным энтропийным вкладом*. Таким образом, изменение стандартной энтропии при реакции (5.7) можно представить в виде

$$\Delta S^0 = S'_{AB} - S'_A - S'_B + R \ln \chi_M = \Delta S' + R \ln \chi_M \qquad (5.11)$$

где χ_M — мольная доля вещества А при его концентрации 1 моль/л. Так как в разбавленных водных растворах молярная концентрация воды равна 55,6, вклад энтропии смешения в общее изменение энтропии приблизительно равен $-R\ln 56 \approx$ ≈ -8 кал · моль$^{-1}$ · град$^{-1}$. При 300 К это соответствует вкладу в свободную энергию, равному $-2{,}4$ ккал · моль$^{-1}$. Если мы выберем другое стандартное состояние или другую единицу измерения концентрации, мы получим, естественно, другой результат. Изменение аддитивного вклада в энтропию ΔS^0_u можно найти из выражения

$$\Delta S^0_u = \Delta S^0 - R \ln \chi_M \qquad (5.12)$$

Дополнение 5.2

ВЫВОД ВЫРАЖЕНИЯ ДЛЯ ПАРЦИАЛЬНОЙ МОЛЯРНОЙ ЭНТРОПИИ

Уравнение (5.10) является обычным выражением для парциальной молярной энтропии. Его легко вывести, исходя из полного дифференциала свободной энергии Гиббса (см. вводный курс физической химии):

$$dG = -SdT + VdP + \Sigma \mu_i dN_i$$

Дифференцируя, получим

$$S_i = (\partial S/\partial N_i)_{T,P,N_{j\neq i}} = -(\partial \mu_i/\partial T)_{P,N_j}$$

В курсе физической химии выводится следующее соотношение:

$$\mu_i = \mu_i^0(T, P) + RT \ln \chi_i$$

где μ_i^0 — стандартный химический потенциал при $\chi_i = 1$. Используя два последних выражения, находим

$$S_i = -(\partial \mu_i/\partial T)_{P,N_j} = -(\partial \mu_i^0/\partial T)_{P,N_j} - R \ln \chi_i = S'_i - R \ln \chi_i$$

Рассмотрим теперь следующие процессы переноса:

Углеводород в неполярном растворителе (мольная доля χ) →
→ Углеводород в воде (мольная доля χ), (5.13)

Чистый жидкий углеводород → Углеводород в воде (мольная доля χ). (5.14)

В процессе (5.13) в отличие от процесса (5.14) вклад энтропии смешения сокращается. Изменение аддитивного вклада в энтропию для процесса (5.14) можно найти путем измерения стандартной энтропии реакции ΔS^0: $\Delta S_u = \Delta S^0 + R \ln \chi$. Получив изменение аддитивной энтропии для второго процесса, мы можем сравнить его непосредственно с изменением энтропии для реакции (5.13).

ДАННЫЕ ПО СВОБОДНОЙ ЭНЕРГИИ ПЕРЕНОСА

Теперь мы можем, наконец, рассмотреть результаты экспериментов по определению свободной энергии переноса. В табл. 5.7 приведены термодинамические параметры, характеризующие перенос различных углеводородов из неполярных растворителей в воду. Приведены аддитивные вклады в свободную энергию и энтропию. Видно, что хотя энергетически растворение в воде выгоднее, чем в неполярных растворителях ($\Delta H < 0$), большое по абсолютной величине отрицательное изменение энтропии при переносе в воду приводит к тому, что изменение полной свободной энергии становится положительным, а процесс переноса невыгодным. Например, при переносе метана из бензола в воду выделяется 2,8 ккал/моль тепла, но сопровождающее этот процесс изменение энтропии на -18 кал $\cdot$ град$^{-1}$ $\cdot$ моль$^{-1}$ увеличивает ΔG_u^0 до $+2{,}6$ ккал $\cdot$ моль$^{-1}$. Сходные результаты получены и для других углеводородов и других неполярных растворителей.

Как интерпретировать эти результаты? Обычно предполагают, что в воде неполярные углеводороды могут вызывать формирование «льдоподобных» клатратов, которые теснее окружают молекулы углеводородов, чем это происходит в кристаллических гидратах (см. рис. 5.15). Это предположение весьма разумно, так как диссоциированные углеводороды способны вызывать повышение температуры замерзания воды. Этот процесс приводит к образованию льдоподобных структур и, следовательно, к формированию водородных связей. Последний процесс является экзотермическим (т.е. происходит с выделением энергии), что объясняет отрицательное значение ΔH при переносе. Однако упорядочение молекул воды в более организованные структуры, естественно, понижает энтропию, что объясняет отрицательное значение ΔS.

Таблица 5.7

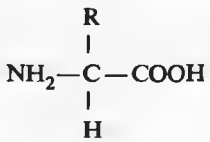
ТЕРМОДИНАМИЧЕСКИЕ ПАРАМЕТРЫ ПЕРЕНОСА РАЗЛИЧНЫХ УГЛЕВОДОРОДОВ ИЗ НЕПОЛЯРНЫХ РАСТВОРИТЕЛЕЙ В ВОДУ ПРИ 25°С[1)]

Реакция переноса	ΔH, ккал $\cdot$ моль$^{-1}$	ΔS_u, кал $\cdot$ град$^{-1}$ $\cdot$ моль$^{-1}$	ΔG_u, ккал $\cdot$ моль$^{-1}$
CH_4 в бензоле → CH_4 в H_2O	$-2{,}8$	-18	$+2{,}6$
CH_4 в эфире → CH_4 в H_2O	$-2{,}4$	-19	$+3{,}3$
CH_4 в CCl_4 → CH_4 в H_2O	$-2{,}5$	-18	$+2{,}9$
C_2H_6 в бензоле → C_2H_6 в H_2O	$-2{,}2$	-20	$+3{,}8$
C_2H_6 в CCl_4 → C_2H_6 в H_2O	$-1{,}7$	-18	$+3{,}7$
Жидкий C_3H_8 → C_3H_8 в H_2O	$-1{,}8$	-23	$+5{,}05$
Жидкий н-C_4H_{10} → н-C_4H_{10} в H_2O	$-1{,}0$	-23	$+5{,}85$
Жидкий C_6H_6 → C_6H_6 в H_2O	$0^{2)}$	$-14^{2)}$	$+4{,}07^{2)}$

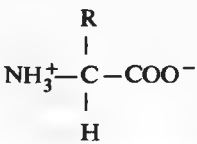
[1)] W. Kauzmann, Adv. Protein. Chem., 14, 1 (1959).
[2)] Измерено при 18° С.

ВЗАИМОДЕЙСТВИЕ БОКОВЫХ ГРУПП НЕПОЛЯРНЫХ АМИНОКИСЛОТ С ВОДОЙ

Результаты, аналогичные приведенным выше, были получены и для аминокислот. Рассмотрим свободную энергию переноса различных аминокислот из этанола в воду (табл. 5.8). При этом процессе молекулы переходят из нейтральной, неполярной формы в этаноле



в цвиттерионную форму



Этот переход вследствие сольватации цвиттерионов полярными молекулами воды дает заметный вклад в ΔG переноса. Важным параметром является *разность* $\Delta G_{\text{пер}}$ между ΔG переноса данной аминокислоты и ΔG в случае глицина. Такая разность может служить мерой вклада в суммарную энергию боковой цепи данной аминокислоты. Из табл. 5.8 ясно, что $\Delta G_{\text{пер}}$ возрастает с увеличением размеров боковой цепи. Весьма неожиданно, что вклад группы $-CH_2$ в $\Delta G_{\text{пер}}$ оказывается одинаков для различных молекул. В табл. 5.8 показано, что разница в ΔG переноса между этаном и метаном почти такая же, как и в случае аланина и глицина. Более того, различия в ΔG между аланином и глицином такие же, как и между лейцином и валином. Подобные результаты дают основание предположить, что отдельные группы $-CH_2$ взаимодействуют с растворителем как более или менее независимые и аддитивные единицы этих небольших простых молекул.

Таблица 5.8

СВОБОДНАЯ ЭНЕРГИЯ ПЕРЕНОСА РАЗЛИЧНЫХ СОЕДИНЕНИЙ ИЗ ЭТАНОЛА В ВОДУ ПРИ 25° С[1)]

Соединение	ΔG, ккал/моль	$\Delta G_{\text{пер}}$, ккал/моль	Соединение	ΔG, ккал/моль	$\Delta G_{\text{пер}}$, ккал/моль
Глицин	−4,63	0	Вклад группы $-CH_2$		
Аланин	−3,90	+0,73	Этан	+3,02	–
Валин	−2,94	+1,69	Метан	+2,26	–
Лейцин	−2,21	+2,42	Этан — метан	–	+0,76
Изолейцин	−1,69	+2,97	Аланин — глицин	–	+0,73
Фенилаланин	−1,98	+2,65	Лейцин — валин	–	+0,73
Пролин	−2,06	+2,60			

[1)] C. Tanford, J. Am. Chem. Soc., **84**, 4240 (1962).
Дополнительные данные см. в табл. 2.3.

ПОНЯТИЕ О ГИДРОФОБНЫХ ВЗАИМОДЕЙСТВИЯХ

Результаты, приведенные в табл. 5.7 и 5.8, показывают, что неполярные боковые группы аминокислот предпочитают находиться в неполярном, неводном окружении, т.е. имеют тенденцию собираться вместе и образовывать кластеры неполярных групп. Силы, стабилизирующие такие кластеры, называют гидрофобными. Результаты рентгеноструктурного анализа многих белков подтверждают, что неполярные боковые группы действительно группируются в неводных, «маслянистых» внутренних областях белка (см. например, рис. 2.32). Отметим, что «движущей силой» гидрофобных взаимодействий является изменение энтропии; изменение же энтальпии, напротив, им противодействует. Следовательно, когда две или большее число первоначально сольватированных гидрофобных групп собираются вместе во внутренней области белка, $\Delta H > 0$. Это означает, что повышение температуры будет смещать равновесие не в сторону диссоциации, а в сторону «гидрофобного связывания»[1)], при условии что диапазон рассматриваемых температур достаточно мал, чтобы величина ΔH оставалась положительной. (Теория гидрофобных эффектов изложена в работе Pratt, Chandler, 1977.)

НАРУШЕНИЕ ГИДРОФОБНЫХ ВЗАИМОДЕЙСТВИЙ МОЧЕВИНОЙ

Приведенные результаты ясно показывают, как велика биологическая роль гидрофобных взаимодействий. Именно они в значительной мере ответственны за стабилизацию белковых молекул с их «спрятанными» в глубине структуры углеводородными боковыми группами. Кроме того, конформация белков должна быть чувствительной к составу растворителя, особенно если изменение этого состава влияет на распределение углеводо-

Таблица 5.9

ИЗМЕНЕНИЕ СВОБОДНОЙ ЭНЕРГИИ ПРИ ПЕРЕНОСЕ РАЗЛИЧНЫХ АМИНОКИСЛОТ ИЗ ВОДЫ В 8М РАСТВОР МОЧЕВИНЫ ПРИ 25° С[1)]

Аминокислота	ΔG, ккал/моль	$\Delta G_{\text{пер}}$, ккал/моль	Аминокислота	ΔG, ккал/моль	$\Delta G_{\text{пер}}$, ккал/моль
Глицин	+0,10	0	Фенилаланин	−0,60	−0,70
Аланин	+0,03	−0,07	Тирозин	−0,63	−0,73
Лейцин	−0,28	−0,38			

[1)] P. L. Whitney, C. Tanford, J. Biol. Chem., **237**, 1735 (1962).

родов между органической фазой и окружающим растворителем. Рассмотрим мочевину — обычный денатурирующий агент для белков. Одно время считали, что денатурирующее действие мочевины на структуру белка связано с конкуренцией ее за образование внутримолекулярных водородных связей, но это представление, по-видимому, неверно, так как тогда вода сама могла бы быть эффективным денатурирующим агентом. Каков же в таком случае механизм действия мочевины на распределение углеводородных белковых цепей между водной и органической фазами?

В табл. 5.9 приведены данные о свободной энергии переноса различных аминокислот из воды в 8М раствор мочевины. Для оценки вклада боковых групп вычислены $\Delta G_{\text{пер}}$, по-

[1)] Это следует из отношения Вант-Гоффа

$$d(\ln K)/dT = \Delta H/RT^2$$

где K — константа равновесия реакции.

лученные вычитанием ΔG для глицина из ΔG для других аминокислот, как это делалось выше. Очевидно, что боковые группы предпочитают воде раствор мочевины, причем степень этого предпочтения коррелирует с размером группы. Таким образом, в присутствии мочевины выигрыш в энергии, в результате которого боковые группы предпочитают находиться внутри белковой молекулы, а не в окружающем ее растворителе, уменьшается на несколько сотен калорий. Поскольку у большинства белков неполярные остатки составляют обычно от 30 до 50%, ясно, что даже в небольшом белке, состоящем из 150 остатков, изменение конформации, вызванное 8 М раствором мочевины, может быть чрезвычайно велико.

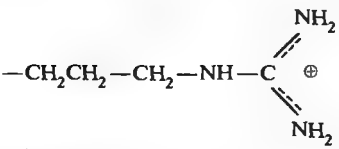
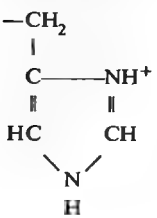
5.8. Ионные взаимодействия

УСЛОВИЯ ВОЗНИКНОВЕНИЯ ИОННЫХ ВЗАИМОДЕЙСТВИЙ В БЕЛКАХ

Многие ионизованные боковые группы в белках способны образовывать солевые мостики, которые играют роль дополнительных невалентных связей, стабилизирующих нативную структуру. В области нейтральных pH карбоксильные группы аспарагиновой и глутаминовой кислот, а также С-концевого остатка заряжены отрицательно. Положительные заряды могут быть локализованы на аминогруппах боковых цепей лизина и N-концевых аминокислот, гуанидиниевых группах боковых цепей аргинина и заряженных имидазольных кольцах гистидина (который ионизуется при нейтральных pH).

Таблица 5.10

НЕКОТОРЫЕ СОЛЕВЫЕ МОСТИКИ, ОБНАРУЖЕННЫЕ В ДЕЗОКСИГЕМОГЛОБИНЕ[1)]

Положительные заряды	Отрицательные заряды	Примечание
Гуанидиниевая группа	Боковая цепь аспарагиновой кислоты	Связывает α_1- и α_2-субъединицы
	$-CH_2-COO^-$	
ε-Аминогруппа лизина	С-концевой карбоксил	Связывает α_2- и β_1- , а также α_1- и β_2-субъединицы
$-CH_2CH_2CH_2CH_2-NH_3^+$	$-COO^-$	
Имидазольное кольцо гистидина	Боковая цепь аспарагиновой кислоты	Внутренняя связь в β-субъединицах
	$-CH_2-COO^-$	

[1)] M. F. Perutz, Nature, **228**, 726 (1970).

Данные рентгеноструктурного анализа подтверждают наличие в белках ионных пар. Некоторые из обнаруженных в дезоксигемоглобине солевых мостиков описаны в табл. 5. 10. В их формировании принимают участие и ионизованные боковые группы, и С-концевые карбоксилы. Примечательно, что ионные взаимодействия нарушаются, когда гемоглобин переходит в оксиформу (см. гл. 17).

УПРОЩЕННОЕ РАССМОТРЕНИЕ ИОННЫХ ВЗАИМОДЕЙСТВИЙ

Оценить вклад ионных взаимодействий в стабилизацию нативной структуры белка довольно трудно. Рассмотрим изменение свободной энергии, связанное с приведением в контакт двух противоположно заряженных ионов (см. Kauzmann, 1959). Определим разность ΔG между свободными энергиями конечного состояния, в котором заряды сближены до расстояния, равного сумме их ионных радиусов (когда они вступают в контакт), и начального состояния, в котором заряды разнесены настолько далеко, что энергия их взаимодействия пренебрежимо мала. Это изменение свободной энергии есть просто работа, которую нужно совершить над системой зарядов, чтобы привести их в соприкосновение. Для зарядов противоположных знаков эта работа отрицательна и равна

$$\Delta G = -\int_{\infty}^{a} (z_1 z_2 e^2/\varepsilon r^2)\, dr = \tag{5.15}$$

$$= z_1 z_2 e^2/\varepsilon a \tag{5.16}$$

где r — переменное расстояние между зарядами, z_1 и z_2 — валентности ионов, e — заряд электрона, ε — диэлектрическая проницаемость. Параметр a — это расстояние, равное сумме ионных радиусов.

Выражение (5.16) носит весьма приближенный характер. В частности, оно выведено в предположении о неизменной диэлектрической проницаемости ε, хотя очень вероятно, что она зависит от r, так как микроскопическое окружение зарядов меняется по мере их сближения. Если предположить, что $z_1 = -z_2 = 1$ и $\varepsilon = 80$ (т.е. равно диэлектрической проницаемости воды), то при $a = 4$ Å $\Delta G \approx -1$ ккал/моль. В случае, когда сближение зарядов происходит вблизи поверхности белка, некоторые электростатические силовые линии проходят через белок, а не через растворитель. Это приводит к уменьшению эффективной диэлектрической проницаемости и тем самым к увеличению (по абсолютной величине) отрицательного изменения свободной энергии. Вместе с тем исходное расстояние между зарядами в развернутой белковой молекуле может быть недостаточно велико, чтобы даже незначительно увеличить энергию электростатических взаимодействий. Изменение свободной энергии при образовании солевых мостиков для такого развернутого состояния будет соответственно меньше.

Изменение энтальпии, связанное с формированием солевых мостиков, можно получить из уравнения (5.16), если предположить, что a не зависит от температуры.

$$\Delta H = -T^2 [d\,(\Delta G/T)/dT] = (z_1 z_2 e^2/\varepsilon a) +$$
$$+ T(z_1 z_2 e^2/\varepsilon^2 a)(d\varepsilon/dT) = (\Delta G/\varepsilon)[d(\varepsilon T)/dT] \tag{5.17}$$

Изменение энтропии определяется следующей формулой:

$$\Delta S = -d\Delta G/dT = (z_1 z_2 e^2/\varepsilon^2 a)(d\varepsilon/dT) =$$
$$= (\Delta G/\varepsilon)(d\varepsilon/dT) \tag{5.18}$$

РОЛЬ ЭНТРОПИИ В ОБРАЗОВАНИИ ИОННЫХ ПАР

Обе производные в уравнениях (5.17) и (5.18) для водных растворов отрицательны. Это означает, что ΔH формирования ионных пар в воде *положительно*, т.е. образование солевых мостиков энергетически невыгодно. Напротив, из того, что ΔS положительно, следует, что движущей силой этого процесса является энтропийный фактор. В согласии с качественными выводами, сделанными на основе такой простой электростатической модели, находятся литературные данные. Например, реакция между Mg^{2+} и ATP^{4-}, приводящая к образованию $MgATP^{2-}$, сопровождается изменением энтропии на $+29{,}4$ кал · град · $^{-1}$ · моль$^{-1}$, благоприятным для формирования комплекса, тогда как $\Delta H = +3{,}3$ ккал/моль, т.е. *препятствует* его образованию.

Конечно, все четыре уравнения (5.15) — (5.18) являются довольно грубым приближением. В частности, растворитель рассматривается как непрерывная среда с неизменной диэлектрической проницаемостью. Не учитывается ни изменение диэлектрической проницаемости в зависимости от расстояния между ионами, ни образование особых комплексов между ионами и молекулами воды. Таким образом, хотя результаты, полученные с помощью этих уравнений, весьма поучительны, не следует забывать, что это только грубая оценка.

Тот факт, что движущей силой образования ионных комплексов являются энтропийные эффекты, был весьма неожиданным. Скорее можно было думать, что они неблагоприятны для этого процесса, так как комплекс безусловно обладает меньшей трансляционной свободой, чем отдельные ионы. Тем не менее полученные результаты можно объяснить на основе простых качественных представлений о воде. Ионы легко сольватируются благодаря сильным ион-дипольным взаимодействиям с молекулами воды. При формировании ионных пар освобождаются высокоупорядоченные и иммобилизованные молекулы воды, что увеличивает «беспорядок» или конфигурационную свободу растворителя. Это приводит к значительному росту общего изменения энтропии, связанного с образованием ионного комплекса. Нарушение энергетически выгодных взаимодействий ионов с водой объясняет, кроме того, положительное изменение энтальпии, связанное с образованием комплекса.

СРАВНЕНИЕ ГИДРОФОБНЫХ И ИОННЫХ ВЗАИМОДЕЙСТВИЙ

Тот факт, что солевые мостики найдены в нативных структурах белков, означает, что они играют определенную роль в стабилизации этих структур. Однако, как уже упоминалось, трудно оценить, насколько велика эта роль. Прочность солевых мостиков должна была бы увеличиваться при снижении дизлектрической проницаемости среды, скажем, при добавлении к воде неполярных растворителей, например диоксана или этилового спирта. Тем не менее при их добавлении белки денатурируют, откуда следует, что расположение углеводородных белковых цепей в неполярном окружении более выгодно, чем усиление электростатической стабилизации, и, следовательно, в нативной белковой структуре гидрофобные взаимодействия важнее электростатических.

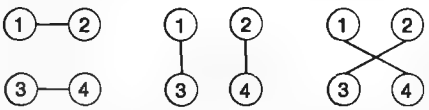
5.9. Дисульфидные связи

Мы рассмотрели три типа невалентных сил, играющих важную роль в формировании белковой структуры: водородные связи, гидрофобные и ионные взаимодействия. Дисульфидные мостики между остатками цистеина относятся к *ковалентным* связям, стабилизирующим белковую конформацию. Такие мостики ограничивают число возможных трехмерных конфигураций, которые может принимать макромолекула. Таким образом, они уменьшают разность в конфигурационной энтропии между нативным и денатуриро-

ванным (при неразрушенных дисульфидных связях) состояниями. Без ограничений, создаваемых дисульфидными связями, денатурированное состояние имело бы значительно бóльшую конфигурационную свободу.

ВОЗМОЖНЫЕ СПОСОБЫ СВЯЗЫВАНИЯ N ПОЛУЦИСТИНОВ

Рассмотрим сначала, каким числом способов (гипотетически) могут сформироваться дисульфидные мостики из четного числа полуцистинов. Обозначим различные полуцистины цифрами 1, 2, 3 ... . Ясно, что два полуцистина могут связаться единственным образом. Для четырех полуцистинов существуют три возможности:



В случае шести полуцистинов у каждого из них имеются пять партнеров для образования пары, а для четырех оставшихся полуцистинов, как мы уже видели, существуют три различных способа, которыми они могут связаться для формирования двух дополнительных мостиков. Таким образом, общее число мыслимых способов, которыми шесть полуцистинов способны образовать три дисульфидных мостика, равно $5 \cdot 3 = 15$. В общем случае при наличии N полуцистинов (N четное) для каждого из них существует $N - 1$ способ образовывать первый дисульфидный мостик, $N - 3$ способов сформировать второй мостик (поскольку два остатка уже связаны) и т.д.; таким образом, для общего числа W гипотетических комбинаций имеем

$$W = (N - 1)(N - 3)(N - 5) \ldots \text{ для четных } N \tag{5.19}$$

(Попробуйте написать выражение для нечетных N.) Совершенно очевидно, что число способов, которыми можно образовать заданное число дисульфидных мостиков, быстро растет с увеличением N.

ОБРАЗОВАНИЕ СПЕЦИФИЧЕСКИХ ДИСУЛЬФИДНЫХ СВЯЗЕЙ

Почему для нативной структуры белка характерна одна и только одна система дисульфидных связей? Этот факт имеет по крайней мере три правдоподобных объяснения.

1. Хотя есть много гипотетических возможностей, только одна из них разрешена геометрически или стерически. Таким образом, белок принимает такую конформацию, в которой может реализоваться именно данное расположение дисульфидных связей.

2. Пары полуцистинов могут образоваться множеством способов, но локализацию дисульфидных связей определяют особые условия биосинтеза белка.

3. Восстановленный белок сворачивается в наиболее стабильную конформацию, и образуется система дисульфидных связей, как бы фиксирующих эту конфигурацию. Другие расположения дисульфидных связей также возможны, но соответствующие им структуры термодинамически менее стабильны.

РЕОКИСЛЕНИЕ РИБОНУКЛЕАЗЫ С ОБРАЗОВАНИЕМ НАТИВНОГО БЕЛКА

Вопросами образования специфических дисульфидных связей занимались Анфинсен и др. (1961 г.), которые выбрали для работы рибонуклеазу. Белок рибонуклеаза (мол. масса около 14000) имеет четыре дисульфидные связи и не содержит свободных групп SH. Со-

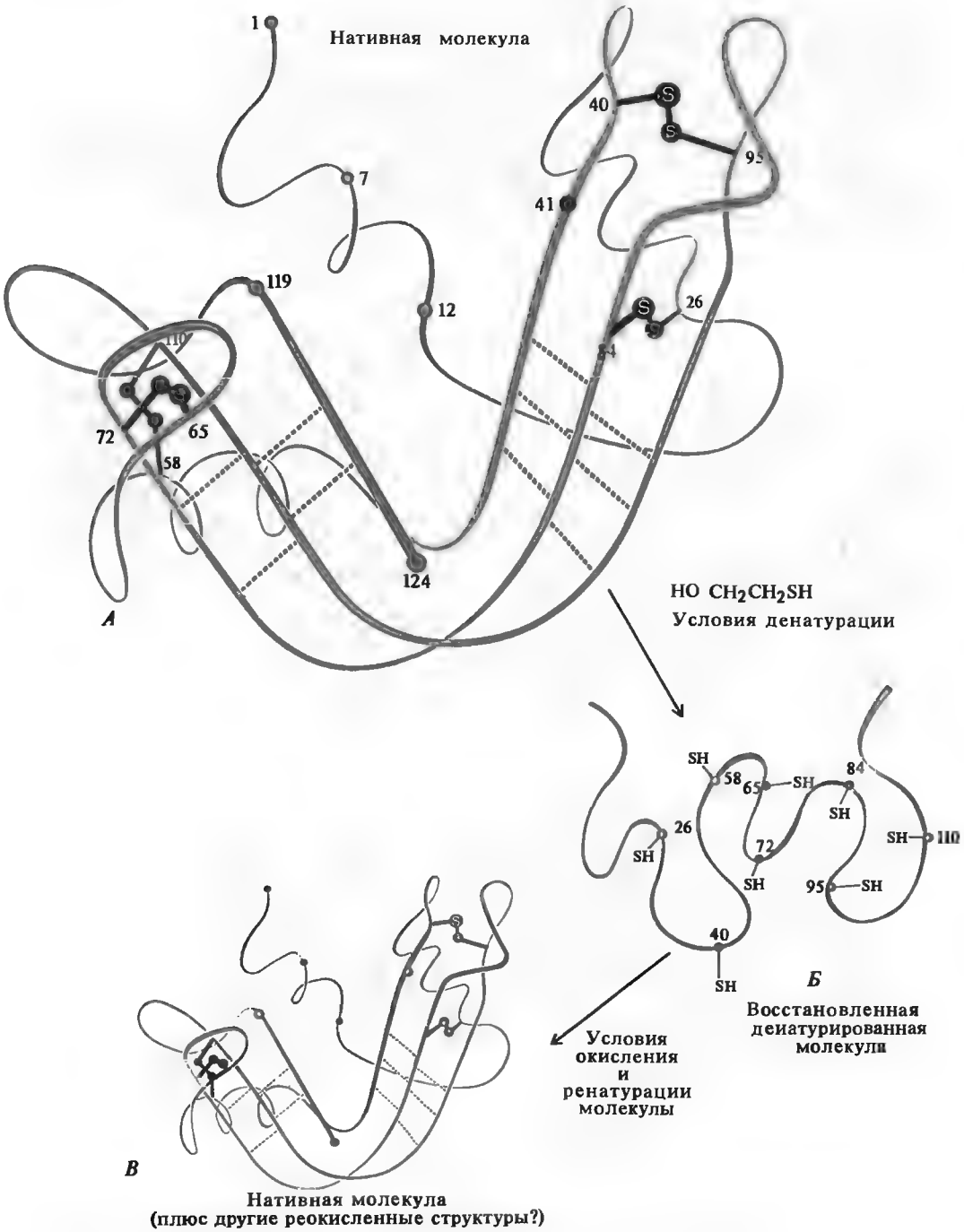


РИС. 5.16. Восстановление и реокисление рибонуклеазы. (Рисунок Ирвинга Гейса.)

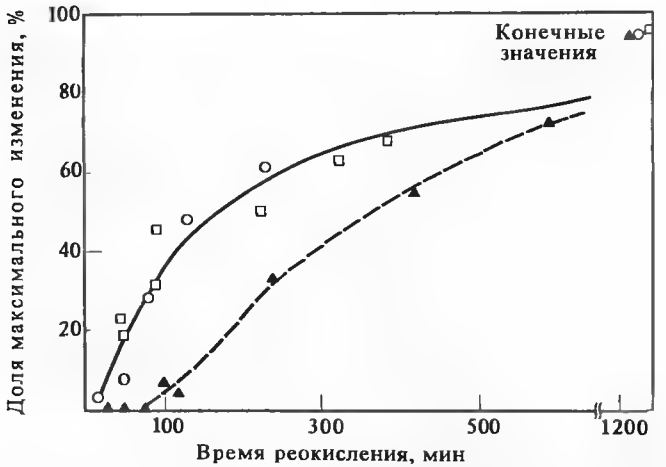


РИС. 5.17. Реокисление восстановленной рибоиуклеазы. Сплошиая кривая — физические свойства; пунктирная кривая — ферментативная активность. Квадратики — оптическое вращение; кружки — титры свободных групп SH; треугольники — ферментативвая активность. [C.B.Anfinsen et al., Proc.Natl.Acad.Sci. USA, **47**, 1309 (1961).]

гласно уравнению (5.19), число способов, которыми четыре дисульфидных мостика могут образоваться из восьми полуцистинов, равно 105.

На рис. 5.16 приведена схема эксперимента, выполненного Анфинсеном и др. Нативную рибонуклеазу сначала денатурировали β-меркаптоэтанолом ($HOCH_2CH_2SH$) в 8М растворе мочевины; при этом происходило восстановление дисульфидных мостиков и молекула прииимала развернутую конформацию. Затем мочевину удаляли и раствор подвергали воздействию воздуха (O_2) при умеренно щелочиых pH (pH ≈ 8) для того, чтобы реокислить полуцистины. Полученные результаты приведены на рис. 5.17, где показаны изменения различных ферментативных свойств в зависимости от времени реокисления. Кружками отмечено уменьшение числа свободных групп SH, квадратиками — изменение оптического вращения при 366 нм. Как видно из рисунка, оба типа точек ложатся на одну кривую. Это дает основание полагать, что уменьшение числа свободных групп SH идет параллельно процессу образования упорядоченной структуры, что сопровождается увеличением оптического вращения.

На рис. 5.17 представлены также даиные по восстановлению ферментативной активности. Кривая имеет отчетливый лаг-период и в начале отстает от кривой изменения физических свойств. В конечном счете, однако, формируются молекулы с нормальной активностью, задержка же в восстановлении активности означает, что молекулы реокисляются ни по принципу «одна за одной», ни по принципу «все или ничего». Зависимость длительности лаг-периода от концентрации белка дает основание предположить, что происходит не только внутримолекулярное, но и межмолекулярное спаривание групп SH. Дисульфидные связи перетасовываются до тех пор, пока не сформируется именно такая схема дисульфидных мостиков, которая свойственна нативному ферменту.

Активные молекулы, образовавшиеся в процессе реокисления, идентичны молекулам нативной рибонуклеазы. Следовательно, мы можем заключить, что второе объяснение причин образования характерных для данной молекулы дисульфидных мостиков не имеет оснований, так как правильная система дисульфидных связей формируется даже in vitro. Справедливость первого из предложенных объяснений проверялась результатами реокис-

ления в 8М растворе мочевины. После полного окисления мочевину удаляли и обнаруживали, что полученные таким образом молекулы сохраняли только 1% активности. Такой низкий уровень активности ясно указывает, что возможно формирование одной или нескольких систем дисульфидных мостиков, отличных от нативной. То, что предсказывается 1% активности, если все 105 гипотетических систем дисульфидных связей могут встречаться с одинаковой вероятностью, а активность наблюдается только для одной из них — нативной, вероятно, следует считать случайным совпадением. Кроме того, при добавлении к «беспорядочно» окисленному веществу SH-реагента перетасовка дисульфидных связей происходила до тех пор, пока не образовывались нативные молекулы.

Эти интересные эксперименты свидетельствуют в пользу третьего из приведенных выше возможных объяснений. Таким образом, мы приходим к заключению, что белковая молекула свертывается с образованием наиболее стабильной конформации, которая затем еще больше стабилизируется путем образования дисульфидных мостиков. Такая точка зрения получает дальнейшее подтверждение при исследовании некоторых белков, не имеющих дисульфидных мостиков, и тем не менее формирующих нативную структуру и способных к обратимой денатурации.

РЕОКИСЛЕНИЕ ПРОИНСУЛИНА С ОБРАЗОВАНИЕМ НАТИВНОГО ИНСУЛИНА

Была исследована способность еще нескольких белков к образованию правильной структуры после восстановления дисульфидных связей. Для всех них кроме инсулина, явно выпадающего из общей картины, были получены сходные результаты. На рис. 1.11 показана структура инсулина; она состоит из А-цепи, содержащей 21 остаток, и В-цепи, содержащей 30 остатков. А- и В-цепи соединены между собой двумя дисульфидными мостиками. Кроме того, в А-цепи имеется мостик между полуцистинами 6 и 11. При денатурации инсулина его цепи перепутываются, и при реокислении не удается получить достаточного количества нативного белка. Следует иметь в виду, однако, что in vivo инсулин синтезируется как белок-предшественник — проинсулин (см. рис. 1.11). Далее эта молекула подвергается ферментативному расщеплению, фрагмент из остатков с 31-го по 63-й удаляется, и получается функционально активный инсулин. При восстановлении и реокислении проинсулина иммунологическая активность, свойственная нативному белку, восстанавливается. Более того, обрабатывая такой проинсулин трипсином, можно получить биологически активный инсулин. Таким образом, дисульфидные связи самопроизвольно формируются в проинсулине и затем сохраняются в инсулине. Без них инсулин не способен принять нативную конформацию. Возникает естественный вопрос: находится ли инсулин в термодинамически наиболее стабильной конформации, по крайней мере в отношении расположения дисульфидных связей?

5.10. Данные о структуре белков

ЧИСЛО ВОЗМОЖНЫХ КОНФОРМАЦИЙ

Выше мы рассмотрели различные факторы, играющие роль в формировании трехмерной структуры нативных молекул белка. Число конформаций, которое может принимать белковая молекула, значительно уменьшается из-за стерических препятствий, возникающих при наложении групп ближайших вдоль аминокислотной цепи остатков. Например, конформационная карта для аланина (см. рис. 5.4) показывает, что разрешенная часть конформационного пространства составляет для этого остатка менее чем 10% всей области в плоскости ϕ — ψ. Однако даже при таком уменьшении свободы выбора мы не можем

(хотя бы приблизительно) определить коиформацию длинной полипептидной цепи. Рассмотрим энергетическую контурную карту для остатка, несущего боковую цепь типа $-CH_2R'$ (см. рис. 5.8). Предположим, что каждая из низкоэнергетических областей (I, II и III) соответствует определенной конформации, так что всего нам нужно рассмотреть три конформации. Тогда для цепи, состоящей из 100 остатков, существует $3^{100} \approx 10^{48}$ энергетически возможных (по крайней мере если речь идет о взаимодействии ближайших соседей) конформаций. Это огромное число. Конечно, наложение атомов, расположенных друг от друга в аминокислотной цепи достаточно далеко, уменьшают его, но не до такой степени, чтобы дать обозримое число возможных конформаций. Поэтому особенно замечательно то, что гидрофобные взаимодействия, водородные связи и солевые мостики способны свести все это великое множество структур к одной-единственной конформации.

СРАВНЕНИЕ СТРУКТУР ЛИЗОЦИМА И МИОГЛОБИНА

Для того чтобы более четко установить, какую роль играют различные взаимодействия в определении пространственной конформации белков, рассмотрим кристаллическую структуру лизоцима и миоглобина. Структура лизоцима схематически изображена на рис. 1.12, 2.31, 2.32 и 2.33, миоглобина — на рис. 1.4,*А*. Эти белки имеют ряд общих черт, весьма важных для формирования наблюдаемых структур. Например, аминокислотные остатки стремятся находиться в разрешенных конформациях (см. рис. 5.12,*А* для лизоцима); полярные группы обычно располагаются на поверхности, а гидрофобные группируются во внутренней области молекулы, недоступной для растворителя.

Несмотря на некоторое сходство, эти два белка имеют совершенно разную конформацию, что еще раз подчеркивает важную роль аминокислотной последовательности и размера цепи в определении конкретной структуры. Рассмотрим сначала лизоцим (129 остатков), морфология которого может быть описана вытянутым эллипсоидом (45 × 30 × 30 Å) с удаленной клинообразной областью. Образовавшаяся щель представляет собой активный центр фермента. Молекула содержит лишь небольшое число участков, подобных α-спирали (остатки 5 — 15, 24 — 34, 88 — 96), и несколько более коротких участков, остатки которых имеют почти такие же углы ϕ и ψ, как и у α-спирали или 3_{10}-спирали [1]. Важная особенность структуры — наличие β-слоя (рис. 5.18). Он образован тремя антипараллельными цепями (остатки 42 — 48, 49 — 54 и 57 — 61), связанными водородными связями, в которых участвуют амидные и карбоксильные группы, а также атомы боковых цепей серина, треонина, аспарагина и глутамина. Некоторые из перечисленных боковых групп принадлежат не остаткам самого β-слоя, а остаткам, расположенным дальше по цепи. Считается, что эта β-структура играет очень важную роль в определении нативной конформации фермента и ее стабилизации. Сомнительно, однако, чтобы указанная область со столь большим числом поперечных водородных связей была так же важна для формирования структуры в целом, как гидрофобные взаимодействия, поскольку сама молекула белка может денатурировать в водном растворе мочевины.

Трехмерная структура миоглобина, состоящего из 153 остатков, значительно отличается от структуры лизоцима, особенно конформацией скелета. Молекула напоминает массивный, сплющенный у полюсов эллипсоид размером 44 × 44 × 25 Å. Ее наиболее замечательная особенность состоит в том, что 121 из 153 остатков находится в

[1] В принятых ранее обозначениях условная запись 3_{10} соответствовала 3 остаткам на виток и 10 атомам в цикле, образующемся при замыкании водородной связи. В этих обозначениях α-спирали соответствует запись $3{,}6_{13}$. В современной системе обозначений запись N_M расшифровывается следующим образом: M есть минимальное число витков, на которых укладывается целое число N остатков. В этой системе 3_{10}-спирали отвечает запись 3_1-спираль, а α-спирали — 18_5-спираль.

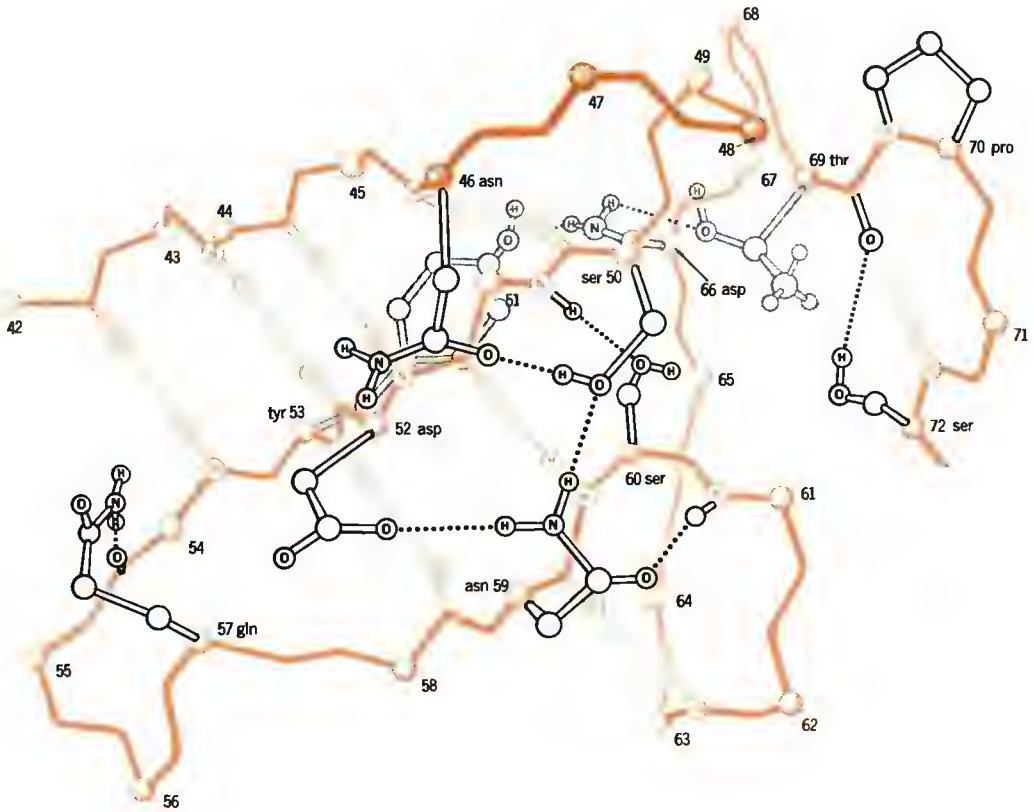


РИС. 5.18. β-Слой в лизоциме. Водородные связи показаны пунктирными линиями. (Рисунок Ирвинга Гейса.)

α-спиральных участках, длина которых колеблется от 7 до 26 остатков. Эти спирали упакованы так, что в целом сформировалась «глобулярная» структура. Схематическое изображение молекулы миоглобина с α-спиральными участками, помеченными буквами от A до H, приведено на рис. 1.4,*A*. В глобуле имеется внутренняя полость, где помещается очень важная в функциональном отношении группа — гем, целиком спрятанная в эту полость (за исключением двух гидрофильных остатков пропионовой кислоты). Атом железа гема связан координационными связями с шестью атомами, каждый из которых отдает электронную пару его орбиталям. Эти шесть атомов расположены по углам октаэдра. Четыре положения в координационном многограннике заняты атомами азота порфириновых колец, пятое — атомом азота имидазольного остатка боковой цепи. Шестое положение занято кислородом в оксимиоглобине и остается свободным в дезоксимиоглобине. Помимо координационных связей с железом положение порфириновых колец стабилизируется гидрофобными взаимодействиями с большим числом гидрофобных остатков, как бы выстилающих полость, в которой эти кольца расположены. В отличие от лизоцима в миоглобине совершенно отсутствуют структуры типа β-слоев.

Многообразие взаимодействий, которые могут иметь место между аминокислотными остатками, создает предпосылки для образования различными аминокислотными после-

довательностями миожества упорядочеиных структур. Очевидио, однако, что аминокислотная последовательность играет первостепенную роль в определении трехмерной структуры белка и общие ограничения, налагаемые стерическими факторами и требованием, чтобы полярные боковые группы располагались на поверхности молекулы, а неполярные внутри нее, не достаточны для того, чтобы установилось однозначное соответствие между последовательностью аминокислотных остатков и той единствеиной конформацией, которую она принимает.

5.11. Предсказание структуры белков

Главной целью исследователей, изучающих структуру белка, является предсказание трехмерной конформации молекулы по ее аминокислотной последовательиостн. Из всего сказанного выше ясно, что это чрезвычайно трудиая задача. Например, для белка, состоящего всего из 100 аминокислотных остатков в одной полипептидиой цепи, конформацию с минимальной свободиой энергией необходимо искать (если преиебречь концевыми эффектами) в 200-мерном пространстве, т.е. одновременно изменять 100 углов ϕ и 100 углов ψ. Даже если мы ограничимся только стерически разрешенными углами ϕ и ψ, число возможных конформаций такого белка будет равно $3^{100} \approx 10^{48}$, так как большинству остатков, обладающих боковой группой (таких, как L-аланин), соответствуют три различные области в плоскости $\{\phi; \psi\}$ (см. рис. 5.8).

Все это лишний раз подчеркивает, с какими огромными числами приходится сталкиваться при решении многомерных задач такого типа. И это только часть проблемы. Анализ конформаций скелета обходит стороной серьезный и весьма спорный вопрос об относительной стабилизации различных конформаций растворителем. Следует также принять во внимание конформации боковых групп и их влияние на стабилизацию конформаций скелета.

МЕТОДЫ ПРЕДСКАЗАНИЯ СТРУКТУРЫ

Непомерная сложность задачи заставила исследователей обратиться к полуэмпирическим методам. Они также ограничились предсказанием некоторых основиых элементов вторичной структуры белков, таких, как α-спирали и β-слои.

Хотя различных методов предсказания структуры довольно много, все они основываются на гипотезе о том, что *в основном* конформация белка определяется локальными, ближними взаимодействиями. Это означает, что короткий полипептид с данной последовательностью с большой вероятностью примет ту же конформацию, что и полипептид с той же последовательностью в крупной молекуле белка.

Существует несколько экспериментальных подтверждений этой гипотезы. Села, Анфинсен и их сотрудники сиитезировали полипептидную «петлю», содержащую остатки 64—82 лизоцима (рис. 5.19, Arnon et al., 1971). Этот синтетический полипептид, будучи введен кроликам и козам, индуцирует у них образование антител, способиых реагировать с нативным лизоцимом и с выделенной из него пептидиой петлей. Вместе с тем, хотя полипептидная петля и нативный лизоцим способны подавлять иммунологические взаимодействия с синтетическими веществами, цептиды, обработанные кислотой (т.е. такие, у которых мостики S—S разорваны), не являются ингибиторами. Эти результаты свидетельствуют о том, что полипептидная петля, по крайней мере до некоторой степени, принимает ту же конформацию, что и петля иативного лизоцима. Другой подход использовали Кабат и др., которые попытались определить коиформацию остатка Y в трипептиде XYZ (см. Kabat, Wu, 1973). Основываясь на имеющихся в их распоряжении данных, исследователи пришли к заключению, что конформация остатка Y в зиачительной мере опреде-

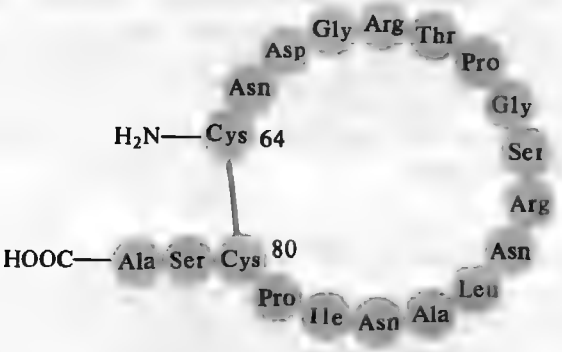


РИС. 5.19. Синтезированная искусственно полипептидная петля, подобная петле лизоцима (остатки 64-82), в которой вместо Ala-76 стоит Cys. [R. Arnon et al., Proc. Natl. Acad.Sci. USA, **68**, 1450 (1971).]

ляется локальными взаимодействиями. Это означает, что конформация Y в последовательности XYZ в неком белке A с большой вероятностью будет такой же, как и в белке B. Их наблюдения были положены в основу одного из методов предсказания структуры белков.

Все эти результаты говорят о том, что можно попытаться предсказывать структуру, не варьируя одновременно все углы ϕ и ψ, а ограничиваясь одновременным рассмотрением только нескольких последовательно расположенных в полипептидной цепи остатков. Основываясь на такой схеме, можно надеяться определить вторичную структуру хотя бы в первом приближении.

ПРИМЕР ПРЕДСКАЗАНИЯ СТРУКТУРЫ

За последние несколько лет было разработано много методов предсказания вторичной структуры. В качестве иллюстрации мы рассмотрим метод, предложенный Чоу и Фасманом (Chou, Fasman, 1974), причем выбор именно этого метода носит до некоторой степени произвольный характер.

Основная идея заключалась в том, чтобы исследовать достаточно большое число структур кристаллических белков с целью определения конформационных «склонностей» различных аминокислотных остатков. Дело в том, что разные остатки обладают неодинаковой вероятностью находиться, например, в составе α-спирали. Данные по синтетическим полипептидам свидетельствуют о том, что некоторые остатки имеют ярко выраженную тенденцию образовывать спиральные полипептиды (см. гл. 20).

Исследование конформационных «склонностей» аминокислотных остатков можно провести на количественном уровне, если рассчитать конформационные параметры для каждого остатка. В табл. 5.11 приводятся результаты такого расчета для 15 белков. Во втором столбце перечислены встречаемости различных остатков в α-спиральных областях ($f_\alpha = n_\alpha/n$, где n — общее число остатков *данного типа* во всех исследуемых белках, n_α — число остатков этого типа в α-спиральных участках). В третьем столбце приводятся конформационные параметры P_α для каждого остатка. P_α определяется следующим образом: $P_\alpha = f_\alpha/\langle f_\alpha \rangle$, где $\langle f_\alpha \rangle$ — среднее значение f_α (т.е. сумма всех f_α, деленная на 20 — общее число различных остатков). Отметим, что если для какого-либо остатка $P_\alpha = 1$, то такой остаток встречается в составе α-спирали с той же частотой, что и в любом другом месте последовательности. Значение $P_\alpha > 1$ свидетельствует о повышенной

Таблица 5.11

ВСТРЕЧАЕМОСТЬ РАЗЛИЧНЫХ АМИНОКИСЛОТНЫХ ОСТАТКОВ В α-СПИРАЛЯХ, β-СЛОЯХ И НЕРЕГУЛЯРНЫХ УЧАСТКАХ ДЛЯ 15 БЕЛКОВ И КОНФОРМАЦИОННЫЕ ПАРАМЕТРЫ P_α, P_β И P_c[1)]

Аминокислота	f_α	P_α	f_β	P_β	f_c	P_c
Ala	0,522	1,45	0,167	0,97	0,311	0,66
Arg	0,282	0,79	0,154	0,90	0,564	1,20
Asn	0,263	0,73	0,113	0,65	0,624	1,33
Asp	0,351	0,98	0,137	0,80	0,514	1,09
Cys	0,278	0,77	0,222	1,30	0,500	1,07
Gln	0,421	1,17	0,211	1,23	0,368	0,79
Glu	0,549	1,53	0,044	0,26	0,407	0,87
Gly	0,190	0,53	0,138	0,81	0,668	1,42
His	0,446	1,24	0,122	0,71	0,432	0,92
Ile	0,358	1,00	0,274	1,60	0,368	0,78
Leu	0,480	1,34	0,209	1,22	0,311	0,66
Lys	0,383	1,07	0,126	0,74	0,491	1,05
Met	0,429	1,20	0,286	1,67	0,286	0,61
Phe	0,402	1,12	0,219	1,28	0,378	0,81
Pro	0,212	0,59	0,106	0,62	0,682	1,45
Ser	0,282	0,79	0,124	0,72	0,594	1,27
Thr	0,295	0,82	0,205	1,20	0,494	1,05
Trp	0,409	1,14	0,203	1,19	0,386	0,82
Tyr	0,220	0,61	0,220	1,29	0,560	1,19
Val	0,409	1,14	0,282	1,65	0,309	0,66

[1)] P. Y. Chou, G. D. Fasman, Biochemistry, 13, 211 (1974).

«склонности» данного остатка встраиваться в α-спираль, а $P_\alpha < 1$ означает, что в среднем он встречается в α-спиралях реже, чем в другом месте. Этот параметр до некоторой степени аналогичен параметру Зимма—Брэгга s, характеризующему рост спирали (он обсуждается в гл. 20). При $s = 1$ 50% бесконечной полипептидной цепи является спиралью.

В табл. 5.11 приведены также параметры f_β, P_β, f_c и P_c, во всех отношениях аналогичные f_α и P_α, но относящиеся к β-слоям и нерегулярным структурам.

Интересно исследовать средние значения этих параметров для α-спиралей и β-слоев тех же 15 белков, которые использовались для расчета табл. 5.11. Средние значения $\langle P_\alpha \rangle$ для остатков, локализованных внутри и на границе α-спиральных участков, приведены на рис. 5.20. Таким образом, $\langle P_\alpha \rangle_{nN}$, $\langle P_\alpha \rangle_N$, $\langle P_\alpha \rangle_i$, $\langle P_\alpha \rangle_C$ и $\langle P_\alpha \rangle_{nC}$ представляют собой средние конформационные параметры для N-конца неспирального участка, N-конца спирали, внутренней области спирали, С-конца спирали и С-конца неспирального участка соответственно. Из рисунка видно, что если мы будем, рассматривая короткие участки полипептидной цепи, постепенно смещаться от центра спирали к ее периферии, то $\langle P_\alpha \rangle$ будет уменьшаться, причем, как только мы пересечем границу спирали, $\langle P_\alpha \rangle$ станет меньше единицы. Это наводит на мысль, что на концах спиральных последовательностей обычно располагаются остатки, имеющие в среднем тенденцию препятствовать образованию спирали и ее росту. Напротив, центральные области спиралей заняты кластерами остатков, предрасположенными к образованию спирали. Эти кластеры могут быть местами инициации спирали, от которых она растет в обоих направлениях.

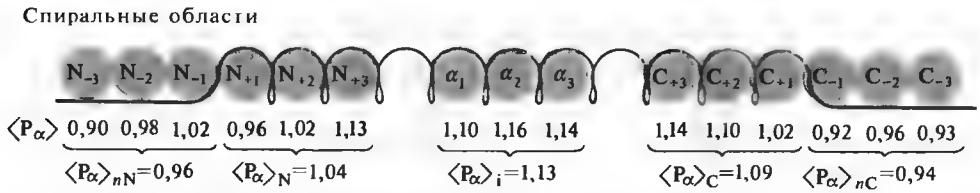


РИС. 5.20. Сравнение величин $\langle P_\alpha \rangle$ для граничных и центральных областей спиралей. [P.Y. Chou, G.D. Fasman, Biochemistry. **13**, 211 (1974).]

Таблица 5.12

КЛАССИФИКАЦИЯ АМИНОКИСЛОТНЫХ ОСТАТКОВ ПО ИХ СКЛОННОСТИ ВСТРАИВАТЬСЯ В α-СПИРАЛЬНЫЕ И β-СТРУКТУРНЫЕ ОБЛАСТИ БЕЛКОВ[1)]

Спиральные остатки	P_α	Классификация	«β-структурные» остатки	P_β	Классификация
$Glu^{(-)}$	1,53	H_α	Met	1,67	H_β
Ala	1,45	H_α	Val	1,65	H_β
Leu	1,34	H_α	Ile	1,60	H_β
$His^{(+)}$	1,24	h_α	Cys	1,30	h_β
Met	1,20	h_α	Tyr	1,29	h_β
Gln	1,17	h_α	Phe	1,28	h_β
Trp	1,14	h_α	Gln	1,23	h_β
Val	1,14	h_α	Leu	1,22	h_β
Phe	1,12	h_α	Thr	1,20	h_β
$Lys^{(+)}$	1,07	I_α	Trp	1,19	h_β
Ile	1,00	I_α	Ala	0,97	I_β
$Asp^{(-)}$	0,98	i_α	$Arg^{(+)}$	0,90	i_β
Thr	0,82	i_α	Gly	0,81	i_β
Ser	0,79	i_α	$Asp^{(-)}$	0,80	i_β
$Arg^{(+)}$	0,79	i_α	$Lys^{(+)}$	0,74	b_β
Cys	0,77	i_α	Ser	0,72	b_β
Asn	0,73	b_α	$His^{(+)}$	0,71	b_β
Tyr	0,61	b_α	Asn	0,65	b_β
Pro	0,59	B_α	Pro	0,62	b_β
Gly	0,53	B_α	$Glu^{(-)}$	0,26	B_β

[1)] P. Y. Chou, G. D. Fasman, Biochemistry, **13**, 222 (1974).

H — остаток, охотно формирующий данную вторичную структуру; h — остаток, склонный к образованию данной вторичной структуры; I — остаток, мало склонный к ее формированию; i — остаток, безразличный к данной вторичной структуре; b — остаток, склонный препятствовать образованию данной вторичной структуры; B — остаток, активно препятствующий формированию вторичной структуры.

Аналогично можно показать, что остатки, для которых $\langle P_\beta \rangle < 1$, обычно располагаются на границах β-слоев, а остатки с относительно высокими значениями $\langle P_\beta \rangle$ чаще всего группируются внутри этих структур.

Эти соображения позволяют построить простую схему предсказания вторичной структуры белков. Суть ее заключается в том, чтобы, приписав конформационные параметры каждому остатку, расположенному в известной последовательности, попытаться затем выделить кластеры остатков полипептидной цепи, склонные образовывать спираль, к которым в определенных точках примыкают остатки, прерывающие формирование этой вторичной структуры. Аналогичную процедуру можно проделать и в отношении β-слоев. Это легко сделать, если в соответствии со значениями конформационных параметров разделить все остатки на следующие классы: формирующие спираль, препятствующие ее образованию, формнрующие β-структуру и препятствующие ее образованию. Такая классификация приводится в табл. 5.12. Используя ее, можно составить набор простых правил для предсказания вторичной структуры белков исходя из их аминокислотной последовательности.

СРАВНЕНИЕ ПРЕДСКАЗАННЫХ И НАБЛЮДАЕМЫХ ВТОРИЧНЫХ СТРУКТУР

По описанной выше схеме была предсказана вторичная структура множества белков, для которых имеются данные, полученные методом рентгеноструктурного анализа. В табл. 5.13 приводится сводка результатов для следующих белков: конканавалин А, гемоглобин миноги, термолизин и ингибитор трипсина. Этих белков нет среди тех 15, для которых были получены основные данные, приведенные в табл. 5.11 и 5.12. В табл. 5.13 перечислены участки последовательности, которые по оценкам должны иметь форму α-спирали и β-слоя, а тажке участки, где эти структуры наблюдаются в действительности. Для ингибитора трипсина — небольшого полипептида, состоящего из 58 остатков, — теоретические и экспериментальные данные прекрасно согласуются между собой. Поскольку это низкомолекулярный белок, локальные взаимодействия, по-видимому, оказывают большее влияние на формирование его структуры, и поэтому можно было ожидать, что полученные для него результаты будут наилучшими. Достаточно хорошее согласие теории и эксперимента получено для других трех белков. В самом деле, 17 из 18 α-спиральных участков и 20 из 22 β-слоев локализованы правильно, и лишь 5 α-спиралей и 10 β-слоев не обнаружены в тех областях, где они должны были бы быть согласно расчетам. Отсюда следует, что результаты предсказания вторичной структуры нельзя объяснить просто случайным совпадением.

Еще один свойственный многим белкам элемент структуры — β-изгиб. Рис. 2.25 показывает, что эта структура дает полипептидной цепи возможность резко менять направление. По аналогии с параметрами P_α и P_β можно рассчитать и конформационные параметры β-изгиба для каждой аминокислоты. Использование их совместно с P_α и P_β позволяет повысить точность конформациоиного анализа. Примечательно, что примерно в 12 белках, недавно исследованных Чоу и Фасманом, около 80% всех остатков находились в составе α-спиралей, β-слоев и β-изгибов и лишь остальные 20% не могли быть отнесены ни к какому определенному типу.

В гл. 2 упоминалось об определении вторичной структуры аденилаткиназы, небольшого полипептида с мол. массой 21 600. На рис. 2.13 сравниваются результаты, полученные для этого белка различными исследователями. Приведены данные для β-изгибов, β-слоев и α-спиралей (сверху вниз). Прямоугольники указывают действительную локализацию этих элементов вторичной структуры в последовательности, а горизонтальные отрезки соответствуют теоретическим данным. Гистограммы в нижней части рисунков дают сводный результат.

Таблица 5.13

СРАВНЕНИЕ ЭКСПЕРИМЕНТАЛЬНЫХ И ТЕОРЕТИЧЕСКИХ ДАННЫХ О ЛОКАЛИЗАЦИИ α-СПИРАЛЕЙ И β-СЛОЕВ В ЧЕТЫРЕХ БЕЛКАХ, НЕ РАССМАТРИВАВШИХСЯ ПРИ РАСЧЕТЕ КОНФОРМАЦИОННЫХ ПАРАМЕТРОВ P_α И P_β[1)]

Белок	α-Спирали		β-Слои	
	рентгеноструктурные данные	теория	рентгеноструктурные данные	теория
Конканавалин А	—	38–43	4–9	3–12
	81–85	81–86	25–29	25–29
	—	155–160	48–55	47–55
	—	180–189	59–66	60–67
			73–78	73–80
			92–97	88–96
			106–116	106–113
			125–132	124–134
			140–144	140–144
			173–177	173–177
			190–199	190–200
			209–215	209–215
			—	229–234
Гемоглобин миноги	12–29	8–24	Нет	
	30-44	—	—	2–6
	45–52	45–52	—	35–43
	62–66	61–78		
	67–88	80–88		
	92–106	92–100		
	—	104–110		
	111–127	115–128		
	132–148	133–147		
Термолизин	—	53–58	4–13	4–17
	65–88	67–74	15–32	20–32
	137–152	137–150	37–46	37–50
	159–180	158–180	52–58	—
	235–246	238–246	60–63	61–66
	259–274	261–271	—	75–84
	280–296	281–259	97–106	98–110
	302–313	301–313	112–116	—
			119–123	120–124
			—	127–131
			—	151–157
			—	192–197
			—	221–225
			—	251–260
			—	272–276
Ингибитор трипсина (из поджелудочной железы)	3–6	2–7	16–24	16–23
	45–56	45–54	27–36	27–38

[1)] P. Y. Chou, G. D. Fasman, Biochemistry, **13**, 222 (1974).

Видно, что сводные расчетные данные находятся в хорошем согласии с данными рентгеноструктурного анализа. Средняя часть рисунка показывает, что в молекуле аденилаткиназы имеется пять β-структурных участков, причем правильно идентифицированы те из них, которые расположены в окрестности остатков с номерами 12, 91 и 116. Эти три участка образуют три центральные цепочки пятицепочечного параллельного β-слоя. Местоположение двух других цепочек установлено на сводной гистограмме не очень точно. Что касается изгибов, то из верхней части рисунка видно, что большинство из них локализованы правильно, хотя некоторые из предсказанных в действительности не встречаются.

В целом все сказанное выше относительно аденилаткиназы иллюстрирует то обстоятельство, что различные полуэмпирические методы позволяют двольно точно предсказывать вторичную структуру белков и что разные подходы к решению этой проблемы могут приводить к сходным результатам.

Предел точности полуэмпирических методов пока трудно установить. Следует также помнить, что точное предсказание вторичной структуры белковой молекулы необходимо, но не достаточно для определения трехмерной конформации. Очень трудно представить себе, как зависит пространственная организация молекулы белка от тех структур, о которых говорят как о «неупорядоченных» (речь идет о структурах, которые нельзя отнести к α-спиралям, β-слоям и т.д.). Неизвестно также, какую роль в формировании трехмерной конформации играют взаимодействия между далеко отстоящими вдоль белковой цепи остатками. Все эти факторы необходимо учесть, прежде чем приступать к предсказанию полной трехмерной структуры белка, исходя из его аминокислотной последовательности.

5.12. Построение молекулярных моделей с помощью ЭВМ

Для предсказания конформации белков и для многих других целей полезно иметь возможность быстро конструировать молекулярные модели белковой структуры. Ясно, что использовать молекулярные модели для построения вручную большого многообразия структур совершенно невозможно. Левинталь и др. (1966 г.) предложили новый, очень остроумный подход к решению этой проблемы, в котором они использовали для построения моделей ЭВМ. Таким методом можно быстро построить молекулярную модель по данным об углах поворота остова молекулы и вывести получившуюся структуру на экран дисплея. Хотя при этом мы видим только двумерную картину, пространственное представление о структуре можно получить, поворачивая модель вокруг какой-нибудь из осей. На рис. 5.21 показаны типичные выведениые на экран дисплея проекции кристаллической структуры миоглобина. Большое впечатление производят наглядность и ясность этих картин. Таким образом, в поисках предпочтительной конформации исследователь может быстро сконструировать и зрительно проанализировать большое количество структур.

Краткие выводы

Структуру белковых молекул определяет множество самых разных, подчас весьма сложных факторов. Длины связей и валентные углы в полипептидной цепи могут считаться по существу фиксированными, а амидные группы, как правило, находятся в плоской *транс*-конформации. Поэтому конформация полипептидного скелета определяется углами внутреннего вращения ϕ и ψ для каждого остатка. Эти углы вращения могут принимать значения в пределах, определенных пространственными и другими взаимодействиями (такими, как амид-амидные дипольные взаимодействия) в цепи. Хотя эти взаимо-

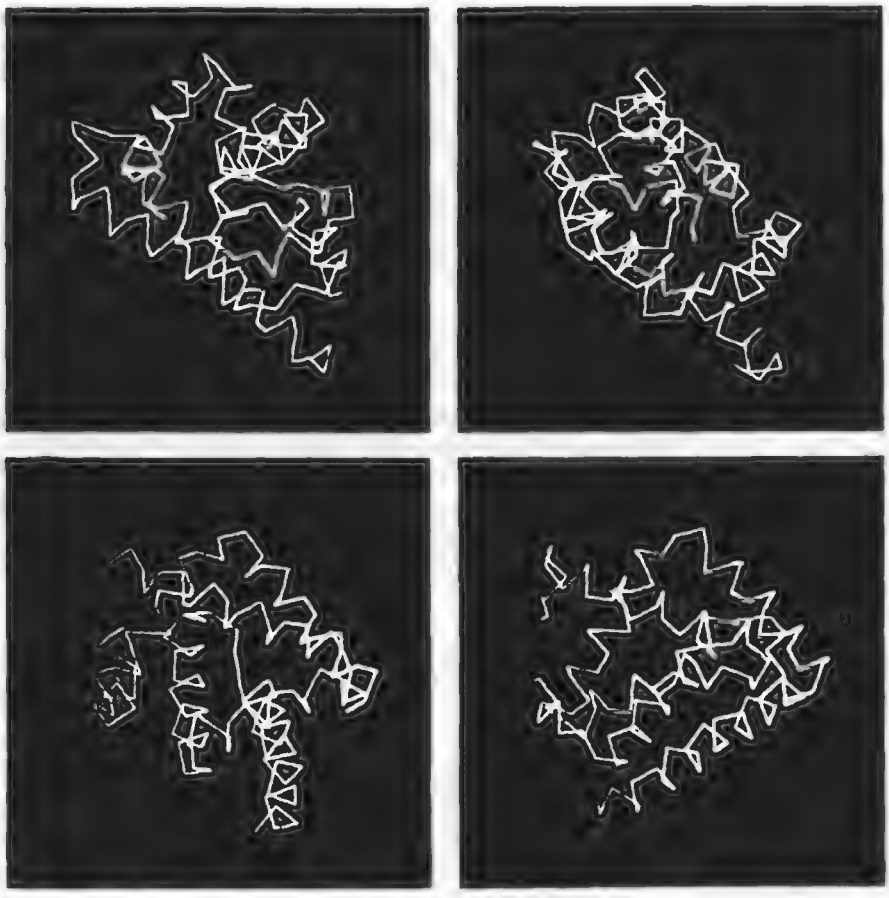


РИС. 5.21. Проекции кристаллической структуры миоглобина на экран дисплея. (Представлена С. Левинталем).

действия исключают значительную часть возможных конформаций, все же их число остается огромным.

Образование водородных связей, дисульфидных мостиков, гидрофобные взаимодействия, обусловленные особыми структурными свойствами воды, а также ионные взаимодействия еще больше ограничивают набор возможных структур полипептида. Совместное влияние всех этих взаимодействий приводит к формированию «нативной» конформации. В определении именно такой конформации ключевую роль играет аминокислотная последовательность белка. Хотя в настоящее время не существует достаточно обоснованного метода предсказания конформации белков по их аминокислотным последовательностям, попытки такого рода делаются и даже иногда оказываются довольно успешными.

Другие вопросы формирования пространственной структуры белковых молекул обсуждаются в гл. 21.

ЗАДАЧИ

5.1. Рассмотрим циклический полипептид с повторяющейся последовательностью XYZX-YZX, состоящий из семи аминокислотных остатков. X, Y и Z — три аминокислотных остатка Ala, Pro и Gly. Следует составить такую последовательность, чтобы цепь замыкалась с наименьшей деформацией. Нужно ответить на вопрос — в каком порядке должны располагаться аминокислотные остатки?

Известный профессор А заявляет, что наилучшее решение такое: Y = Pro, Z = Gly, X = Ala. Робкий студент утверждает, что решение X = Ala, Y = Gly и Z = Pro лучше. Третий участник, философ, ничего не знающий об аминокислотах, говорит, что ни одно из этих решений не является действительно наилучшим. Кто из них прав и почему?

5.2. Белок претерпевает небольшие конформационные изменения, наиболее значительное из которых — перемещение трех плотно упакованных аминокислотных боковых цепей — лейцина, аланина и фенилаланина — из этанолоподобного окружения в белке во внешнюю область, где они погружены в растворитель. Предполагая, что это единственная термодинамически значимая конформационная перестройка, и используя результаты, приведенные в данной главе, рассчитайте константу равновесия этого процесса в воде и в 8 М растворе мочевины при 25°С. Используйте данные этой главы.

5.3. Две части одной белковой молекулы взаимодействуют друг с другом по некоторому характерному закону. В конформации А они находятся на расстоянии r_2 друг от друга, в конформации В — на расстоянии r_1, причем $r_2 > r_1$. Сила взаимодействия $F = -\gamma/r^2$, где $\gamma = aT$, a — постоянная, не зависящая от температуры. Предполагая, что всеми другими взаимодействиями можно пренебречь, рассчитайте константу равновесия, ΔH и ΔS (как функции r_1, r_2, γ и T) для этой конформационной перестройки.

5.4. Экспериментатор измерил свободную энергию переноса ($\Delta G^0_{\text{сумм}}$) неполярного соединения А (мольная доля $\chi = 0{,}1$) из этанола в воду (мольная доля $\chi = 0{,}1$). Он также измерил свободную энергию переноса $\Delta G^{0*}_{\text{сумм}}$ из чистой жидкости А в воду (до мольной доли $\chi = 0{,}1$). На основании экспериментальных результатов он пришел к выводу, что $\Delta G^{0*}_{\text{сумм}} - \Delta G^0_{\text{сумм}} = -RT \ln 56$, и решил, что это прекрасный ответ. Согласны вы с ним или нет? Почему?

5.5. Два полипептидных волокна 1 и 2 имеют точно одинаковую длину, толщину и ширину. Одно из них состоит в основном из длинных цепей полиаланина, другое — из длинных цепей полиглицина. Если эти волокна не растянуты, полипептиды находятся в состоянии беспорядочных клубков, вообще говоря соответствующих конформационным энергетическим картам отдельных остатков. Был обнаружен интересный факт: волокно 1 можно обратимо растянуть до гораздо большей длины, чем волокно 2. На основании этого наблюдения и данных о предпочтительных конформациях остатков глицила и аланила определите, какое из волокон состоит из полиглициновых цепей, а какое — из полиаланиновых. Кратко, но точно обоснуйте сделанный выбор.

ЛИТЕРАТУРА

Общая

Chou P. Y., Fasman G. D., 1978. Empirical predictions of protein conformation, Ann. Rev. Biochem., **47**, 251. [Краткая сводка эмпирических методов.]

Dickerson R. E., Geis J., 1969. The Structure and Action of Proteins, New York, Harper and Row. [Включает прекрасные иллюстрации белковых структур. Особая ценность книги в том, что она дает хорошее зрительное представление о конформациях белков].

Flory P. J., 1969. Statistical Mechanics of Chain Molecules, New York, Interscience. [Имеется перевод: Флори П. Статистическая механика цепных молекул. — М.: Мир, 1971. В гл. 7 приводится подробная сводка энергетических конформационных расчетов полипептидов. Как в большинстве литературных источников того времени, значения углов $\phi = 0°, \psi = 0°$ приписываются конформации, в которой все связи полипептидной цепи находятся в *транс*-положении, что, в соответствии с соглашением, принятым в IUPAC—IUB в 1970 г., отвечает углам $\phi = 180°, \psi = 180°$.]

IUPAC—IUB Commission on Biochemical Nomenclature, 1970. Abbreviations and symbols for the description of the conformation of polypeptide chains, Biochemistry, **9**, 3471. [Сводка определений углов внутреннего вращения в полипептиде и связаниого с этим материала. Публикуется одновременно в J. Biol. Chem., **245**, 6489 and J. Mol. Biol., **52**, 1.]

Kauzmann W., 1959. Some factors in the interpretation of protein denaturation, Adv. Protein Chem., **14**, 1. [Прекрасная статья, посвященная гидрофобным эффектам; обсуждаются также ионные взаимодействия.]

Schultz G., Schirmer R. H., 1979. Principles of Protein Structure, New York, Springer-Verlag. [Всестороннее обсуждение и большая библиография.]

Специальная

Anfinsen C. B., Haber E., Sela M., White F. H. Jr., 1961. The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain, Proc. Natl. Acad. Sci., **47**, 1309.

Anfinsen C. B., Scheraga H. A., 1975. Experimental and theoretical aspects of protein folding, Adv. Protein Chem., **29**, 205.

Arnon R., Maron E., Sela M., Anfinsen C. B., 1971. Antibodies reactive with native lysozyme elicited by a completely synthetic antigen, Proc. Natl. Acad. Sci., **68**, 1450.

Brant D. A., Flory P. J., 1965. The configuration of random polypeptide chains, II: Theory, J. Am. Chem. Soc., **87**, 2791.

Brant D. A., Miller W. G., Flory P. J., 1967. Conformational energy estimates for statistically coiling polypeptide chains, J. Mol. Biol., **23**, 47.

Chou P. Y., Fasman G. D., 1974. Conformational parameters for amino acids in helical, β-sheet and random coil regions calculated from proteins, Biochemistry, **13**, 211.

Chou P. Y., 1974. Prediction of protein conformation, Biochemistry, **13**, 222.

Doty P., 1960. Optical rotation and the structure of polypeptides and proteins. In: Fourth International Congress of Biochemistry, Vol. VIII, Proteins, New York, Pergamon Press, p. 8.

Frank H. S., Wen W.-Y., 1957. Structural aspects of ion—solvent interaction in aqueous solutions: A suggested picture of water structure, Disc. Faraday Soc., **24**, 133.

Gurney R. W., 1953. Ionic Processes in Solution, New York, McGraw-Hill. [Обсуждаются понятия вкладов аддитивной энтропии и энтропии смешения в свободную энергию; см., например, стр. 89 — 92 и следующие разделы.]

Hagler A. T., Honig B., 1978. On the formation of protein tertiary structure on a computer, Proc. Natl. Acad. Sci., **75**, 554.

Kabat E. A., Wu T. T., 1973. The influence of nearest-neighbor amino acids on the conformation of the middle amino acid in proteins: Comparison of predicted and experimental determination of β-sheets in concanavalin A, Proc. Natl. Acad. Sci., **70**, 1473.

Klotz I. M., Franzen J. S., 1962. Hydrogen bonds between model peptide groups in solution, J. Am. Chem. Soc., **84**, 3461.

Levinthal C., 1966. Molecular model building by computer, Sci. American, **214** (6), 42.

Levitt M., Chothia C., 1976. Structural patterns in globular proteins, Nature, **261**, 552.

Némethy G., Scheraga H. A., 1965. Theoretical determination of sterically allowed conformations of a polypeptide chain by a computer, Biopolymers, **3**, 155.

Pratt L. R., Chandler D., 1977. Theory of the hydrophobic effect, J. Chem. Phys., **67**, 3683.

Ramachandran G. N., Ramakrishnan C., Sasisekharan V., 1963. Stereochemistry of polypeptide chain configurations, J. Mol. Biol., **7**, 95.

Глава 6 Конформационный анализ и силы, определяющие структуру нуклеиновых кислот

6.1. Общая характеристика структуры нуклеиновых кислот

Физические и химические свойства нуклеиновых кислот существенно отличаются от свойств белков и полипептидов. Это является следствием совершенно разного химического состава и строения двух указанных классов молекул. В то время как полипептидный остов электрически нейтрален и к нему присоединены боковые цепи приблизительно двадцати типов, остов нуклеиновой кислоты представляет собой сильно заряженный полиэлектролит, который несет боковые группы только четырех (в большинстве случаев) типов. Далее, боковые цепи нуклеиновых кислот проявляют специфическую комплементарность (спаривание оснований), которая отсутствует у аминокислот. Эта комплементарность частично ответственна за образование спиральных палочкообразных структур как в двух-, так и в одноцепочечных молекулах. Кроме того, заряженный остов затрудняет переход нуклеиновых кислот в компактные глобулярные конформации, столь типичные для белков.

Мы рассмотрим наиболее важные факторы, определяющие конформацию нуклеиновых кислот. Прежде всего мы проанализируем ограничения, налагаемые на возможные конфигурации геометрическими и стерическими требованиями. Сразу же станет ясно, что остов нуклеиновой кислоты устроен значительно сложнее полипептидного остова. Затем мы займемся вопросами специфического спаривания оснований и неспецифического стэкинг-взаимодействия, которые играют главную роль в стабилизации упорядоченных форм одно- и двухцепочечных молекул.

6.2. Геометрические аспекты

На рис. 6.1 схематически изображена полинуклеотидная цепь. Повторяющаяся единица цепи содержит шесть скелетных связей и боковую группу — основание, присоединенное к атому $C^{1'}$ рибозы. На рис. 3.1 приведены структурные формулы обычных и многих минорных оснований и указаны системы нумерации атомов для пуриновых и пиримидиновых нуклеозидов. При сопоставлении полинуклеотидного остова с полипептидным, изображенным иа рис. 5.1, становится очевидным, насколько полинуклеотидная цепь сложнее по своему строению, чем полипептидная.

В табл. 6.1 приведены структурные параметры остова для кристаллов полирибоадениловой кислоты; указаны длины связей и некоторые углы между ними. Значения этих параметров фиксированы довольно жестко и примерно одинаковы для разных РНК.

Для того чтобы рассмотреть множество допустимых конформаций цепи, целесообразно остановиться сначала на конформационных характеристиках индивидуальных мононуклеозидных единиц (состоящих из сахара и основания), а затем перейти к рассмотрению стерических ограничений в остове.

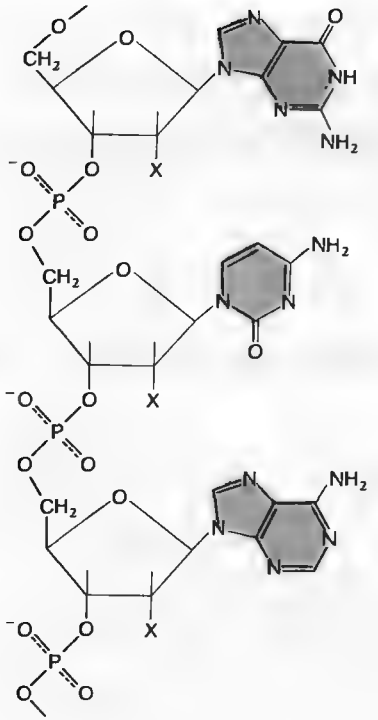


РИС.6.1. Схематическое изображение полинуклеотидной цепи. Для РНК X — это OH, для ДНК — H.

Таблица 6.1

ГЕОМЕТРИЯ ПОЛИРИБОНУКЛЕОТИДНОЙ ЦЕПИ[1,2)]

Связь	Длина связи, Å	Связи	Угол между связями, град
		Геометрия остова	
$C^{4'}—C^{3'}$	1,52	$C^{4'}—C^{3'}—O^{3'}$	110
$C^{3'}—O^{3'}$	1,47	$C^{3'}—O^{3'}—P$	119
$O^{3'}—P$	1,56	$O^{3'}—P—O^{5'}$	104
$P—O^{5'}$	1,57	$P—O^{5'}—C^{5'}$	121
$O^{5'}—C^{5'}$	1,46	$O^{5'}—C^{5'}—C^{4'}$	112
$C^{5'}—C^{4'}$	1,54	$C^{5'}—C^{4'}—C^{3'}$	112
		Геометрия гликозидной связи	
$O^{1'}—C^{1'}$	1,45	$O^{1'}—C^{1'}—N^{9}$	112
$C^{2'}—C^{1'}$	1,54	$C^{2'}—C^{1'}—N^{9}$	106
$C^{1'}—N^{9}$	1,47	$C^{1'}—N^{9}—C^{4}$	127
$N^{9}—C^{4}$	1,35	$C^{1'}—N^{9}—C^{8}$	127
$N^{9}—C^{8}$	1,33		

1) A. Rich et al., J. Mol. Biol., 3, 71 (1961).

2) Приведены данные для молекулы полирибоадениловой кислоты.

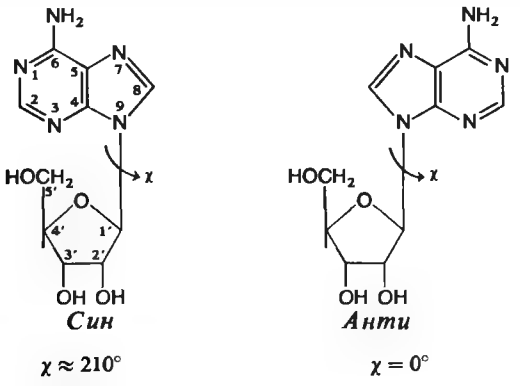


РИС. 6.2. Конформеры аденозина.

6.3. Поворотная изомерия относительно гликозидной связи; конформация рибозы

Основание присоединяется к полинуклеотидному остову с помощью гликозидной связи. Ориентация его по отношению к рибозе описывается углом поворота χ вокруг этой связи в направлении по часовой стрелке. Этому углу приписывается значение 0° для *цис*-конформации связей $O^{1'}$ — $C^{1'}$ и N^9 — C^8. В случае пиримидинов угол поворота определяется ориентацией связей $O^{1'}$ — $C^{1'}$ и N^1 — C^6.

Особый интерес представляют два поворотных измера, обозначаемые как *син* и *анти* (рнс. 6.2). Для *анти*-конформации угол χ равен примерно 0°, а для *син* - около 210°. Известна также форма с углом χ, равным 120°. Насколько одно изомерное состояние оказывается предпочтительнее другого — зависит от конформации рибозы, которая обсуждается в следующем разделе. Вообще говоря, расчеты показывают — и рентгеноструктурные данные, полученные на мононуклеозидах и мононуклеотидах, подтверждают это, — что пиримидины предпочитают конфигурацию *анти,* тогда как пурины способны принимать обе формы. В двойной спирали ДНК и в спиральной двухцепочечной полиадениловой кислоте основания находятся в *анти*-конфигурации.

Как отмечалось выше, ориентация основания относительно гликозидной связи зависит от конформации рибозного кольца. Конформация этого кольца, как правило, такова, что четыре атома находятся в одной плоскости, а пятый расположен выше или ниже этой плоскости. Атомом, выходящим из плоскости, является $C^{2'}$ или $C^{3'}$, а смещение обычно обозначается как *эндо* или *экзо* в зависимости от того, направлено оно в ту же сторону, что и у атома $C^{5'}$, или в противоположную соответственно (см. рис. 3.15). Различные конформации рибозы обозначаются, таким образом, как $C^{2'}$-*эндо,* $C^{2'}$-*экзо,* $C^{3'}$-*эндо,* $C^{3'}$-*экзо*. Все эти конформации приведены на рис. 3.15.

6.4. Углы вращения остова и стерические ограничения

Анализ предпочтительных конформаций остова нуклеиновой кислоты представляет чрезвычайно трудную задачу. Должно быть учтено множество стерических взаимодействий при изменении шести углов вращения в каждой повторяющейся единице, а также исследованы возможности взаимодействия между соседними единицами. Таким образом,

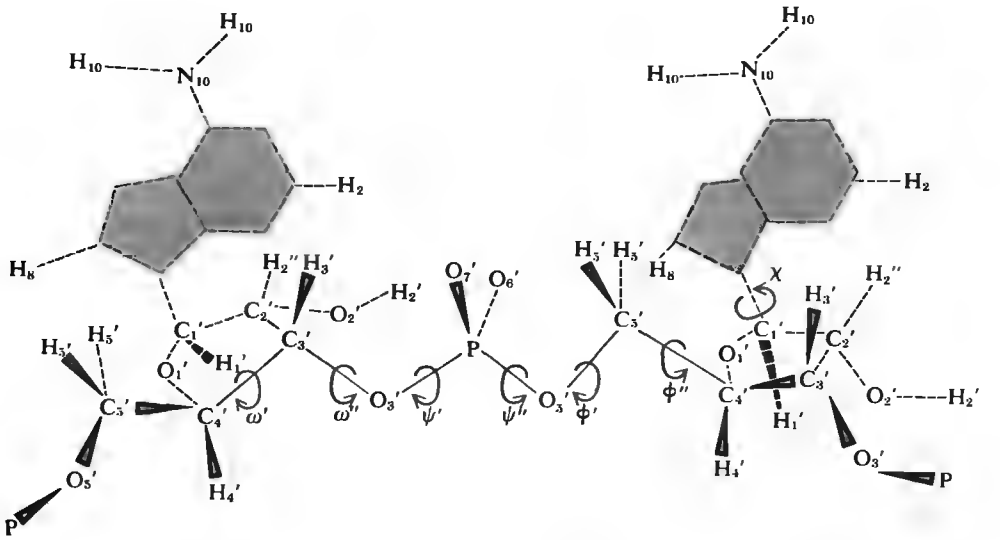


РИС. 6.3. Схематическое изображение цепи полирибоаденилловой кислоты с указанием углов вращения. [W.K. Olson, P.J. Flory, Biopolymers, **11**, 1 (1972).]

сформулированная проблема представляется значительно более сложной, чем анализ предпочтительных конформаций полипептидной цепи.

Для решения этой проблемы мы будет следовать процедуре, предложенной Олсон и Флорн (1972 г.), существенно упрощающей проблему. На рис. 6.3 схематически изображен сегмент полирибоаденилловой цепи; углы вращения остова вокруг отдельных связей обозначены греческими буквами с одним или двумя штрихами.

Весьма существенно, что периодическое расположение пентозных колец обеспечивает независимость вращений вокруг пяти связей остова, следующих за кольцом, от вращений вокруг пяти связей, предшествующих кольцу. Поэтому, если мы выберем повторяющуюся единицу, начинающуюся от одного атома $C^{4'}$ и кончающуюся следующим (см. рис. 6.3), конформации каждой повторяющейся единицы будут независимы друг от друга, и нам останется рассмотреть лишь стерические возможности совместного изменения шести углов вращения (на рис. 6.3 обозначенных ω', ω'', ψ', ψ'', ϕ', ϕ'') в пределах одной повторяющейся единицы.

ВЗАИМОДЕЙСТВИЯ ПЕРВОГО ПОРЯДКА

Систематически исследуя стерические ограничения, возникающие при изменении каждого угла вращения по отдельности, мы значительно упростим проблему. Будем считать, что данная часть остова находится в *цис*-конфигурации, если угол вращения для соответствующей связи равен нулю. Прежде всего проанализируем взаимодействия между теми атомами, для которых взаимные расстояния зависят только от одного угла вращения. Рассмотрим, например, стерические ограничения, которые возникают при изменении единственного угла ψ'. К ним относятся ограничения, связанные с взаимодействием между атомом $C^{3'}$, с одной стороны, и атомами $C^{5'}$, $O^{6'}$ и $O^{7'}$ — с другой; в эти случаях расстояние между $C^{3'}$ и атомами кислорода зависит только от угла ψ'. В то же время взаимодействие между атомами $C^{3'}$ и $C^{5'}$ не попадает в эту группу, поскольку расстояние

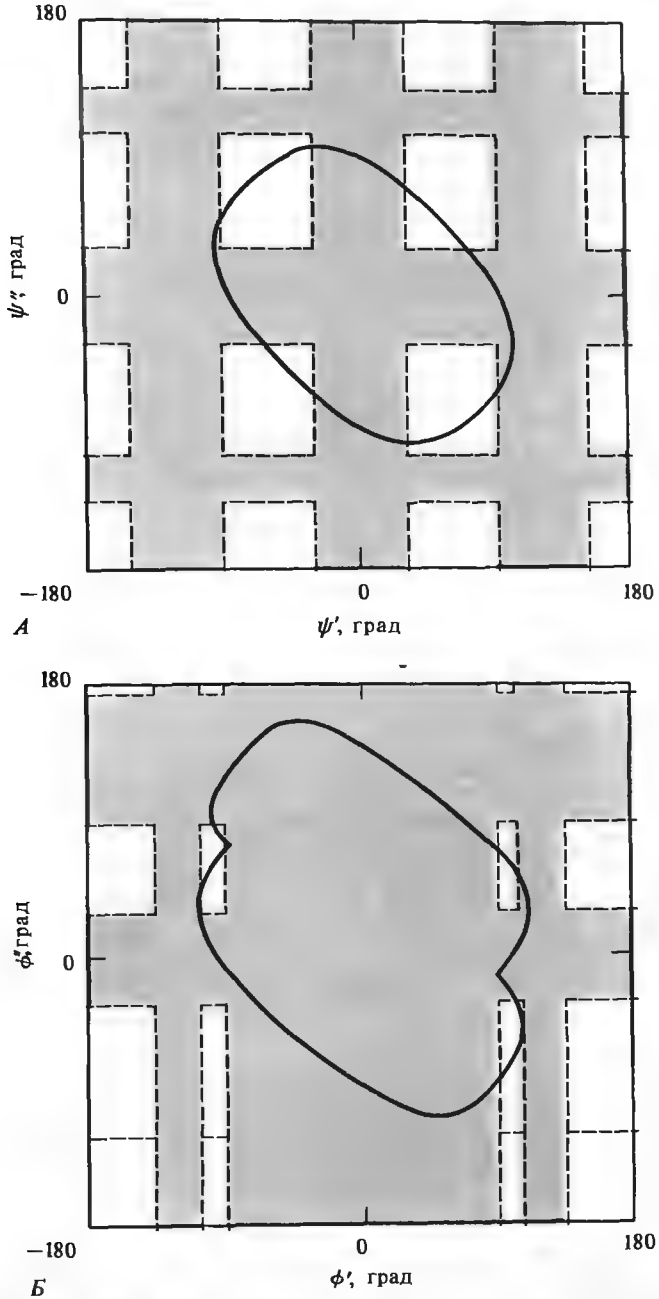


РИС. 6.4. Карты допустимых значений углов вращения. Области значений, запрещенных из-за взаимодействий первого порядка, закрашены, а из-за взаимодействий второго порядка обведены сплошной линией. *А*. Карта для углов ψ'' и ψ'. *Б*. Карта для углов ϕ'' и ϕ'. Небольшая область при $\phi''=-120°$ (горизонтальная пунктирная линия) тоже должна быть исключена. [W.K. Olson, P.J. Flory, Biopolymers, **11**, 1 (1972).]

между ними зависит как от ψ', так и от ψ''. Учет взаимодействий первого порядка позволяет значительно сузить диапазон допустимых значений ψ'; эти значения лежат в интервалах от $-90°$ до $-30°$, от 30 до 90° и в районе $180° \pm 30°$.

Аналогичным образом рассматриваются взаимодействия, зависящие только от ψ''. Карта допустимых значений углов ψ' и ψ'' приведена на рис. 6.4,*А*. Закрашенные участки соответствуют областям значений, стерически запрещенным лишь из-за взаимодействий первого порядка, т.е. взаимодействий, определяемых единственным углом вращения (ψ' или ψ'').

ВЗАИМОДЕЙСТВИЯ ВТОРОГО ПОРЯДКА

Проанализируем теперь, какая часть области значений углов ψ' и ψ'', разрешенных по взаимодействиям первого порядка, исключается из-за взаимодействий второго порядка, представляющих собой взаимодействия между теми атомами, расстояние между которыми зависит одновременно от ψ' и ψ'' (таково, например, расстояние $C^{3'} — C^{5'}$). Взаимодействия второго порядка исключают из области допустимых значений площадь, обведенную сплошной линией на рис. 6.4,*А*. Значительная часть этой площади закрашена, и, следовательно, соответствующие углы уже запрещены из-за взаимодействий первого порядка. Тем не менее небольшая часть «разрешенной» взаимодействиями первого порядка «площади» исключается из-за взаимодействий второго порядка. Оставшаяся незакрашенной площадь (вне области, обведенной сплошной линией) свободна от стерических ограничений второго порядка.

ДАЛЬНЕЙШИЙ АНАЛИЗ

Результаты подобного рассмотрения для углов вращения ϕ' и ϕ'' суммированы на рис. 6.4,*Б*. Здесь, как и ранее, незакрашенные участки отвечают областям значений, разрешенных по взаимодействиям первого порядка, а сплошная линия ограничивает область значений, запрещенных из-за отсутствия места для размещения тех атомов, расстояние между которыми зависит одновременно от ϕ' и ϕ''. В случае углов ω' и ω'' разрешенная область значений ω'' зависит в какой-то степени от угла ω'. Значения последнего ограничены интервалом приблизительно в 15° с центром при 83° для $C^{2'}$-*экзо*- и $C^{3'}$-*эндо*-конформаций и примерно таким же интервалом с центром при 153° для $C^{2'}$-*эндо*- и $C^{3'}$-*экзо*-конформеров. Допустимые значения ω'' лежат в окрестности -145 и $-95°$ в интервале от 10 до 20° в зависимости от ω'.

Действуя указанным способом, можно убедиться, что значительная часть пространственных конфигураций запрещена вследствие взаимодействий первого и второго порядков. Однако необходимо еще рассмотреть взаимодействия второго порядка с учетом других пар двугранных углов (например, ψ' и ω'') и возможные взаимодействия атомов, расстояние между которыми зависит от трех или более углов вращения. Кроме того, мы еще не рассматривали возможность взаимодействий с участием атомов оснований — взаимодействия основание—основание и основание—остов.

Все эти проблемы, однако, упрощаются, поскольку каждое из указанных взаимодействий происходит в тех областях конфигурационного пространства, которые уже были исключены из-за взаимодействий первого и второго порядков. На рис. 6.5 светлые участки отвечают тем областям, где отсутствуют взаимодействия первого порядка при данных ϕ' и ψ''. Сплошная линия ограничивает область значений этих углов, запрещенную из-за взаимодействий второго порядка. Однако эта область уже и так исключается, и, следова-

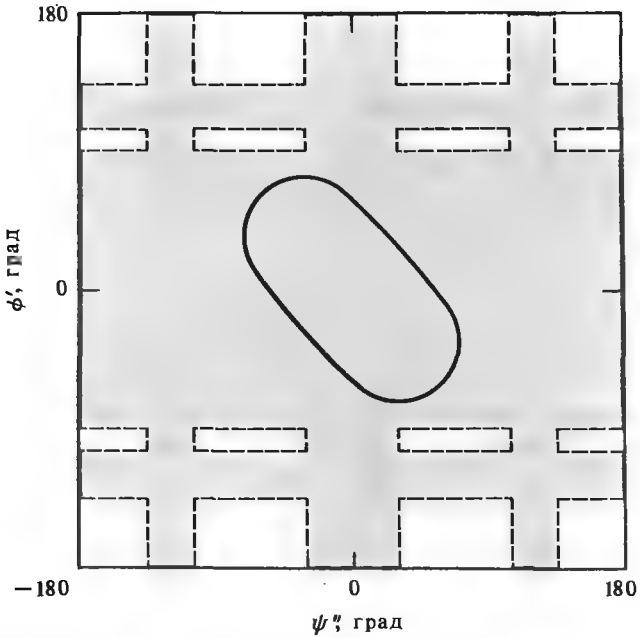


РИС. 6.5. Карта допустимых значений углов ϕ' и ψ''; обозначения те же, что и на рис. 6.4. [W.K. Olson, P.J. Flory, Biopolymers, 11, 1 (1972).]

тельно, взаимодействия второго порядка не сужают область значений, разрешенных взаимодействиями первого порядка.

То же самое справедливо для взаимодействий второго порядка при изменении углов ψ' и ω''. Более того, если значение ω' фиксировано при 83 или 153°, не возникает никаких стерических затруднений при вращении вокруг двух связей, примыкающих к связи $C^{4'} - C^{3'}$ (ω'' и ϕ''). Это обеспечивает независимость вращений внутри одной структурной единицы, расположенной между двумя атомами $C^{4'}$, от вращения в соседних единицах. И наконец, когда углы вращения остова находятся в областях, разрешенных в рамках проведенного выше рассмотрения, отсутствуют стерические ограничения для взаимодействий между соседними основаниями, нет перекрывания оснований с остовом и не возникает помех при одновременном изменении трех или более двугранных углов остова.

Границы стерически разрешенных областей могут быть уточнены таким же способом, как и в случае полипептидов, т.е. с использованием метода межатомных потенциалов. Здесь, однако, вычисления становятся более сложными и должны включать тщательный анализ электростатических взаимодействий между различными группами. При таком анализе приписывается дробный заряд соответствующим атомам фосфатных групп, оснований и рибозы. Расчеты, проведенные этим методом, показывают, что конформация остова смещается в те стерически разрешенные области, которые соответствуют минимальной электростатической энергии, и что в этих областях предпочтительными являются более вытянутые конформации.

Конфигурации, допустимые для остова 2′-дезоксирибонуклеиновой кислоты, лишь незначительно отличаются от таковых для РНК. Различие обусловлено главным образом изменением в энергии, связанным с вращением вокруг связи $C^{3'} - O^{3'}$.

6.5. Силы, стабилизирующие упорядоченные конформации

Хотя расчеты методом межатомных потенциалов позволяют найти те области конфигурационного пространства, в которых молекула может существовать, эти вычисления, как и в случае белков, не могут объяснить образование высокоупорядоченных структур, присущих нуклеиновым кислотам при физиологических условиях. Это связано с тем, что разрешенная стерическими и простыми электростатическими ограничениями область пространства хотя и относительно ограничена, еще достаточно велика, чтобы мог существовать широкий спектр неупорядоченных структур. Более того, исходя из предыдущих вычислений, нельзя понять механизм образования постоянно встречающихся двух- и даже трехцепочечных структур: для этого необходимо рассматривать взаимодействия между цепями. Следовательно, нам нужно вернуться к анализу других типов сил, которые играют первостепенную роль в стабилизации регулярных структур нуклеиновых кислот.

На рис. 3.14 схематически изображена B-форма двойной спирали ДНК в соответствии со структурой, предложенной Дж. Уотсоном и Ф. Криком. Две правые полинуклеотидные спирали переплетены так, что основания расположены внутри структуры, а фосфодиэфирный остов снаружи. Основания образуют прямой угол с осью спирали и уложены в стопку друг над другом; расстояние между соседними основаниями составляет 3,4 Å. На виток спирали приходится десять пар оснований, и, следовательно, шаг спирали равен 34 Å. Цепи ориентированы в противоположных направлениях (т.е. имеют противоположную полярность), так что на любом конце спирали одна цепь несет свободный 5′-фосфат, а другая — свободный 3′-гидроксил.

Еще до того, как была установлена эта структура, было показано, что число пуриновых и пиримидиновых оснований в ДНК одинаково и, более того, число остатков аденина (A) равно числу остатков тимина (T), а число остатков гуанина (G) равно числу остатков цитозина (C). Структура Уотсона—Крика дает объяснение этому сделанному ранее наблюдению. Основания противоположных цепей занимают строго фиксированные положения, что дает им возможность образовывать водородные связи. Этот процесс строго специфичен и приводит к тому, что A может взаимодействовать только к T, а G — с C, поэтому цепи являются взаимно комплементарными. Таким образом, если, двигаясь от 5′- к 3′-концу одной цепи, мы встречаем последовательность AGCTAGGT ... , то в противоположной цепи мы найдем последовательность (прочитанную от 3′- к 5′-концу) TCGATCCA (принято, однако, нумерацию в последовательности вести от 5′- к 3′-концу, так что последние восемь букв следует записать как ... ACCTAGCT). Это комплементарное спаривание оснований между цепями оказывается возможным потому, что ферментативный синтез происходит путем специфического достраивания на одной цепи второй, комплементарной к ней, а не воспроизведения последовательности оснований в этой цепи. Следует подчеркнуть, что, как было показано, в различных экспериментальных условиях структура ДНК хорошо соответствует уотсон-криковской, наблюдаемой в волокнах.

Две важные особенности не были объяснены в рамках проведенного нами конформационного анализа — специфическое спаривание оснований и поразительная параллельность в расположении уложенных в стопку оснований. Эти особенности характерны и для других структур нуклеиновых кислот, в том числе тех из них, которые образуются в случае одноцепочечных полимеров, таких, как тРНК. Они вызывают особый интерес, и мы рассмотрим их более детально.

6.6. Спаривание оснований

Возможны различные типы спаривания оснований, и это многообразие удается наблюдать на опыте.

На рис. 3.11 показано специфическое спаривание оснований по схеме Уотсона и Крика, согласно которой между аденином и тимином (или урацилом) образуются две водородные связи, а между гуанином и цитозином — три. Хотелось бы понять, почему образуются пары оснований именно с этим типом специфичности и, далее, допуская эту специфичность, почему в AT- и GC-парах реализуется именно это конкретное расположение водородных связей. Ответ на эти вопросы, без сомнения, нетривиален.

Хорошо известно, что основания могут ассоциировать, образуя комплексы с помощью водородных связей. Множество таких комплексов из одинаковых оснований обнаружено при кристаллографических исследованиях. Рис. 6.6 иллюстрирует образование водородных связей между основаниями аденина в кристаллах. Такие же ассоциаты из одинаковых элементв наблюдались в кристаллах и других оснований.

При совместной кристаллизации обычных мономерных производных аденина и урацила тоже наблюдается образование пар A — U, однако они *никогда не являются уотсон-криковскими.* Обнаруженные типы комплексов иллюстрирует рис. 6.7. На рис. 6.7,*А* показана уотсон-криковская схема спаривания дезоксиаденозина и тимидина; на рис. 6.7,*Б* — схема связывания, обнаруженного К. Хугстеном (1963 г.) в комплексе 9-метиладенина и 1-метилтимина. В этом комплексе роль акцептора водородной связи играет азот N^7 аденинового кольца (имидазольной части кольца). Эта структура известна как *хугстеновская,* или *имидазольная.* Рис. 6.7,*В* иллюстрирует обратную имидазольную структуру, в которой акцептором водородной связи является атом O^2 урацила вместо O^4. Расчет методом молекулярных орбиталей, выполненный Пульманом и его коллегами (1966 г.), дает для пары аденин — тимнн следующую последовательность структур в порядке убывания их стабильности: имидазольная структура, обратная имидазольная структура, уотсон-криковская структура.

Для GC-пар, которые также исследовались в кристаллах, наблюдается другая ситуация. В этом случае имеет место только уотсон-криковское спаривание, что, вероятно, обусловлено наличием трех водородных связей в этой структуре.

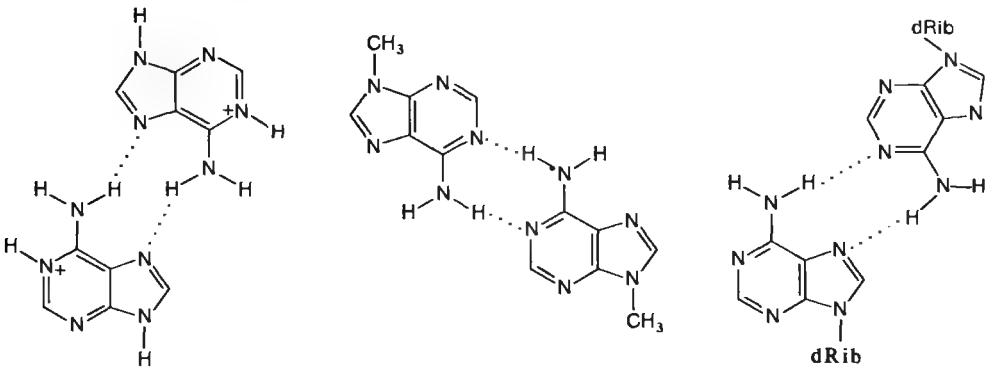
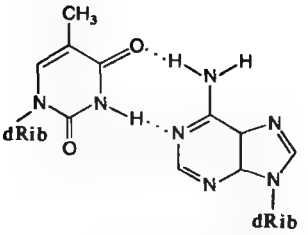
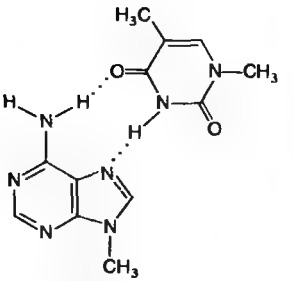


РИС. 6.6. Па́ры, образующиеся между остатками аденина в некоторых кристаллических структурах. [D. Voet, A. Rich, Progr. Nucleic Acid Res.Mol.Biol., 10, 183 (1970).]



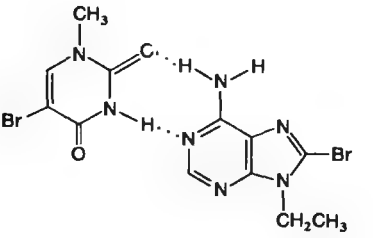
Уотсон-криковская (предполагаемая)

А



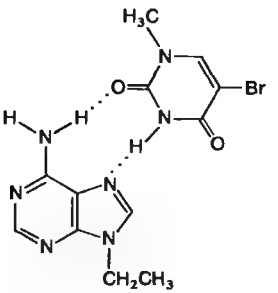
Хугстеновская

Б



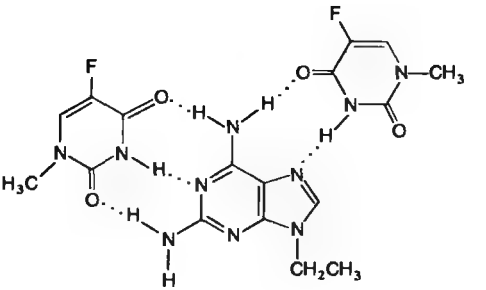
Обратная уотсон-криковская

В



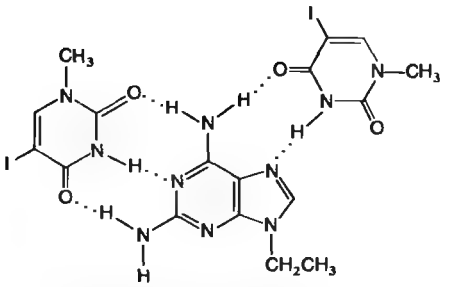
Обратная хугстеновская

Г



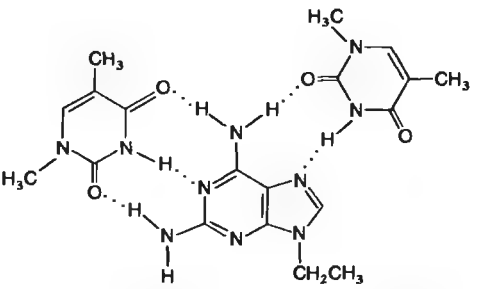
Уотсон-криковская и хугстеновская

Д



Обратная уотсон-криковская и хугстеновская

Е



Уотсон-криковская и обратная хугстеновская

Ж

РИС. 6.7. Различные пары, образующиеся с помощью водородных связей между производными аденина и урацила. [D. Voet, A. Rich, Progr. Nucleic Acid.Res. Mol.Biol., **10**, 183 (1970).]

УОТСОН-КРИКОВСКОЕ СПАРИВАНИЕ В КРИСТАЛЛАХ ДИНУКЛЕОТИДОВ

Хотя уотсон-криковское связывание обычных мономерных производных A и U в кристаллах не наблюдалось, динуклеотид ApU кристаллизуется как мини-спираль, в которой происходит комплементарное уотсон-криковское спаривание оснований. Комплементарные комплексы ApU были исследованы с атомным разрешением А. Ричем и его коллегами (1973 г.); их структура приведена в верхней части рис. 6.8. Если посмотреть на соседние остатки рибозы в сахарофосфатной цепи, то станет ясно, что комплекс закручен вправо. В

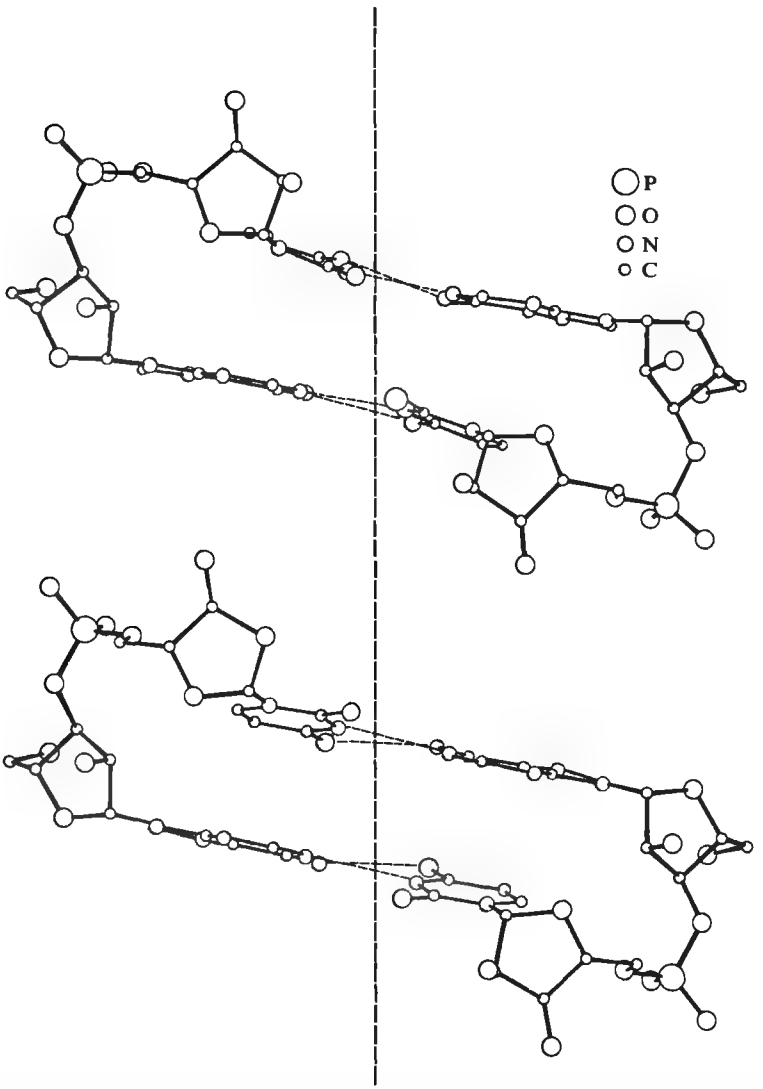


РИС. 6.8. Структура ApU (*вверху*) и спирали РНК с периодом в 11 пар (*внизу*); вертикальная пунктирная прямая — ось спиралей. [J.M. Rosenberg et al., Nature, 243, 150 (1973).]

нижней части рисунка изображен сегмент спирали РНК с 11 парами на виток, в которой реализуется уотсон-криковское спаривание. Такой тип спирали, обнаруженный, например, в вирусных РНК, очень сходен со спиралью ApU. Следовательно, хотя A (или U) могут спариваться сами с собой и хотя в кристаллах смеси мономерных производных A и U обычно образуются комплексы A — U, стабилизированные водородными связями по схеме, отличной от уотсон-криковской, двойная спираль Уотсона—Крика образуется уже из простых динуклеотидов и, как показывают исследования, из poly(A) и poly(U). По-видимому, в случае олигомеров и полимеров играют роль факторы среды и геометрические эффекты, не действующие на мономеры. Они делают уотсон-криковское спаривание более предпочтительным.

СПЕЦИФИЧЕСКОЕ СПАРИВАНИЕ ОСНОВАНИЙ В РАСТВОРЕ

Выше мы рассмотрели образование различных пар оснований в кристаллах. Описанные эксперименты иллюстрируют многообразие взаимодействий, которые могут при этом реализоваться, но ничего не говорят об относительной стабильности различных типов комплексов. Интересно, в частности, определить, будет ли в смеси осиований проявляться тенденция к образованию специфических комплексов или же будет получаться широкое распределение комплексов различных типов. Подобного рода вопросы можно исследовать путем анализа взаимодействий между основаниямн спектроскопическимн методами, такими, как инфракрасная спектроскопия или ЯМР. Ясно, что исследования должны проводиться в растворителе, который является слабым донором или акцептором водородиых связей, чтобы ассоциация нуклеозидов была более четко выраженной.

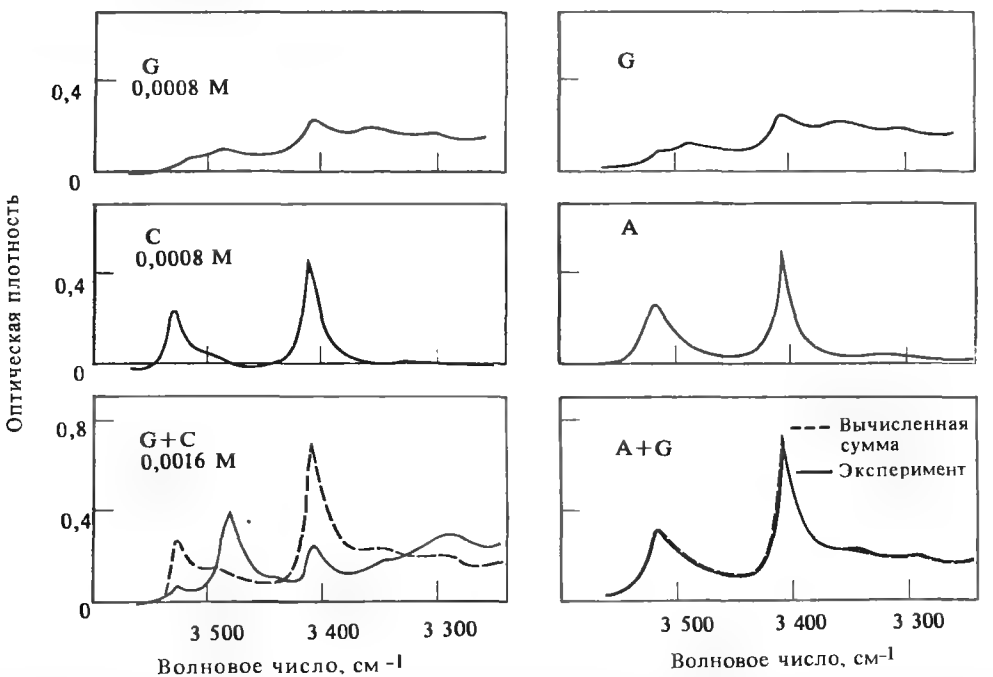


РИС.6.9. Инфракрасные спектры производных гуанина, цитозина и аденина и их смесей, лежащие в области частот колебаний NH-группы. [R.C. Lord, G.J. Thomas, Jr., Devel. Appl. Spectroscopy, 6, 179 (1968).]

Таблица 6.2

КАЧЕСТВЕННЫЕ ХАРАКТЕРИСТИКИ СПЕЦИФИЧНОСТИ СПАРИВАНИЯ ОСНОВАНИЙ, ПОЛУЧЕННЫЕ С ПОМОЩЬЮ ИНФРАКРАСНОЙ СПЕКТРОСКОПИИ[1,2)]

Первое основание	Второе основание				
	C	I	G	T	U
A	O	Cp	O	C	C
U	O	O	O	O	
T	O	O	O		
G	C	O			
I	C				

[1)] Y. Kyogoku et al., Biochim. Biophys. Acta, 179, 10 (1969).
[2)] Относительная стабильность ассоциатов обозначена следующим образом: C — сильное взаимодействие, Cp — среднее, O — взаимодействие отсутствует.

Таблица 6.3

ТЕРМОДИНАМИЧЕСКИЕ ПАРАМЕТРЫ ОБРАЗОВАНИЯ ПАР ОСНОВАНИЙ В ДЕЙТЕРОХЛОРОФОРМЕ ПРИ 25°C[1)]

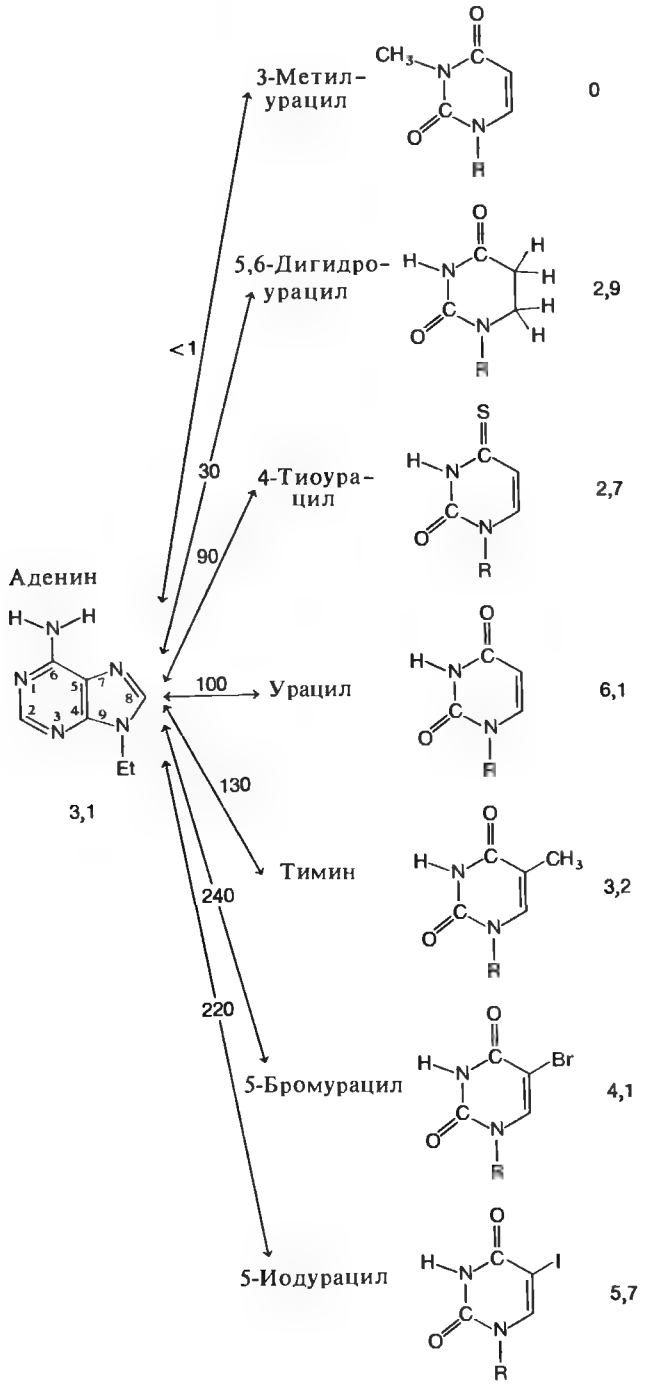
Пара оснований	K, M^{-1}	$-\Delta H$, ккал · $моль^{-1}$	$-\Delta S$, кал · $град^{-1}$ · $моль^{-1}$
U — U	6,1	4,3	11,0
A — A	3,1	4,0	11,4
A — U	100	6,2	11,8
C — C	28	6,3	15
G — G	10^3 — 10^4	(8,5 — 10)[2)]	(15)[3)]
G — C	10^4 — 10^5	(10 — 11,5)[2)]	(15)[3)]

[1)] Y. Kyogoku et al., Biochim. Biophys. Acta, 179, 10(1969).
[2)] Значение получено в предположении, что $\Delta S = -15$ кал × × $град^{-1}$ · $моль^{-1}$.
[3)] Предположительное значение.

Исследования такого рода были выполнены Р. Лордом и А. Ричем и их коллегами. На рис. 6.9 приведены инфракрасные спектры в области частот колебаний N—H-связи для производных гуанина, цитозина и аденина, а также для попарных смесей этих оснований в дейтерохлороформе. Сплошные кривые на левых графиках — наблюдемые спектры для G и C по отдельности при концентрации оснований 0,0008 M и спектр их эквимолярной смеси (суммарная концентрация в этом случае составляла 0,0016 M). Если бы никакого взаимодействия не было, наблюдаемый спектр должен был бы совпадать с суммой двух индивидуальных спектров — пунктирной кривой на нижнем левом рисунке. Однако он явно отличается от этой кривой, что свидетельствует о наличии взаимодействия между G и C. С другой стороны, соответствующие кривые на правом нижнем графике ясно указывают на то, что между G и A взаимодействие отсутствует.

В табл. 6.2 суммированы данные исследований, аналогичных приведенным на рис. 6.8, в которых было проверено наличие попарного взаимодействия для четырех главных и одного минорного оснований (инозина, I). Эти данные отражают результаты качественных наблюдений инфракрасных спектров в смесях, обычно содержащих каждое из оснований в концентрации 0,004 M. Они указывают на высокую спицифичность спаривания оснований, поскольку для многих из исследованных смесей отсутствуют какие-либо признаки взаимодействия. В то же время наблюдается специфическое спариваение A с T (или U) и G с C. Кроме того, обнаружены прочные ассоциаты I—C и более слабые комплексы A—I.

Для AU-пар имеются другие данные, указывающие на образование как уотсон-криковских, так и имидазольных (и/или обратных имидазольных) структур. Тем не менее, не обращая внимания на детали образованных структур, можно сказать, что специфичность взаимодействий оснований друг с другом, наблюдаемая для мономеров, в точности соответствует специфичности, постулированной Уотсоном и Криком для двойной спирали ДНК.


3-Метил-урацил
0
5,6-Дигидро-урацил
2,9
4-Тиоурацил
2,7
Урацил
6,1
Тимин
3,2
5-Бромурацил
4,1
5-Иодурацил
5,7
Аденин
3,1
Et
<1
30
90
100
130
240
220

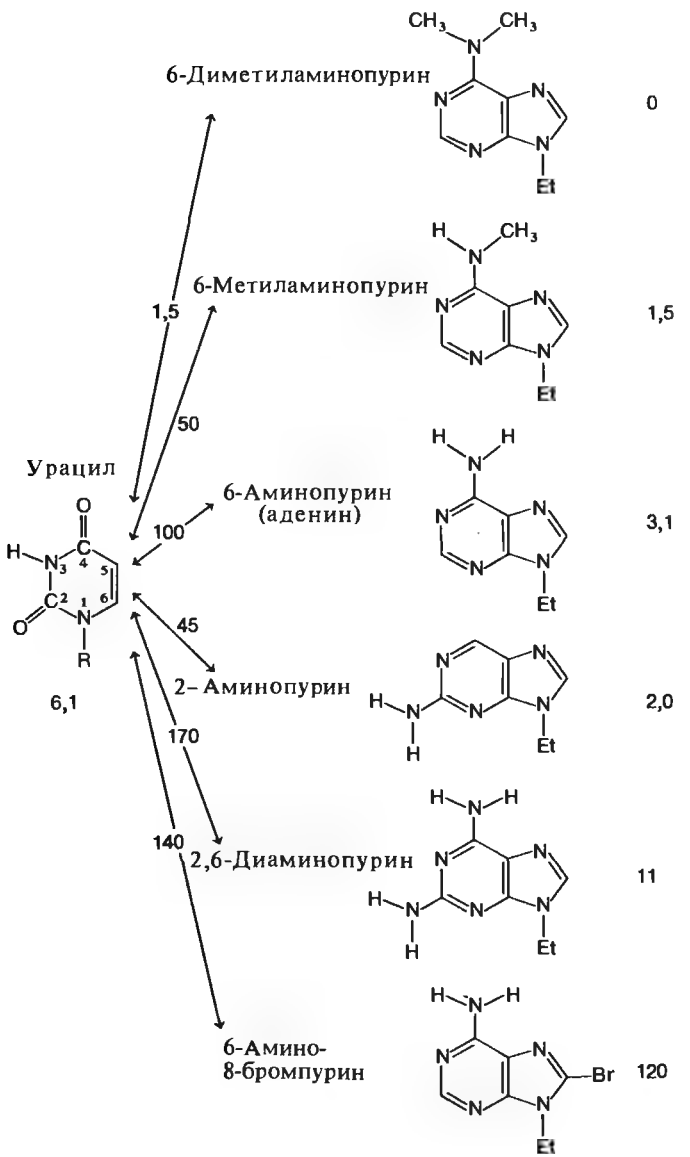


РИС. 6.10. Взаимодействие между производными 9-этиладенина и 1-циклогексилурацила. Растворитель — $CDCl_3$; 25 °C. [R.C. Lord, G.J. Thomas, Jr., Devel, Appl.Spectroscopy, 6, 179 (1968).]

В табл. 6.3 приведены некоторые термодинамические данные о спаривании оснований при 25°С в дейтерохлороформе. Видно, например, что константы связывания при образовании пар оснований из одинаковых моиомерных единиц и уотсон-криковских пар различаются не менее чем в 15 раз, а иногда их отношение достигает 1000 и более. Такое предпочтение уотсон-криковских партнеров обусловлено выигрышем в энергии комплексов, что следует из значительиых различий в энтальпии связывания ΔH для различных пар оснований.

Для дальнейшего выяснения природы тех факторов, которые определяют прочность и специфичнось данного комплекса, рассмотрим ряды производных оснований, которые слегка отличаются от исходных оснований. Такие ряды для A и U приведены на рис. 6.10. Вдоль стрелок, соответствующих взаимодействию между 9-этиладенином и 1-циклогексилурацилом, а также между их производными, и рядом с каждым их этих производных указаны соответствующие константы ассоциации (в M^{-1}). Эти данные позволяют понять детали реакции образования пар. Например, неспособность 3-метилурацила связываться с аденином обусловлена, по-видимому, невозможностью образования циклического димера. Большое значение константы ассоциации 5-бромурацила с аденином объясняется, вероятно, существенным увеличением кислотности NH-группы, вызванным оттягиваним электрона атомом брома; эффект увеличения кислотности NH-группы должен с избытком компенсировать противоположное влияние уменьшения основности кислорода. (Увеличение кислотности NH-группы означает, что водород слабее притягивается к азоту и, следовательно, легче вступает во взаимодействие с другой электроотрицательной группой. Уменьшение основности кислорода означает понижение его сродства к иону водорода.) Можете ли вы обосновать некоторые другие из приведенных результатов?

Здесь мы должны подчеркнуть, что не совсем ясно, почему специфичность столь четко выражена на мононуклеозидном уровне. Создается впечатление, что хотя геометрическая комплементарность характерна для многих пар оснований, ее недостаточно для образования прочного комплекса, т.е. специфические AU- и GC-пары в дополнение к геометрической обладают «электронной» комплементарностью, и этот факт несомненно требует квантовомеханического объяснения. Причина того, что такое объяснение до сих пор не получено, связана с приближенным характером квантовомеханических расчетов; точность этих расчетов недостаточна, чтобы объяснить относительно небольшую разницу в энергии между «специфическими» и «неспецифическими» комплексами.

Результаты описанных выше экспериментов приводят к выводу, что спицифичность спаривания A — U и G — C является неотъемлемым свойством самих мономерных оснований — свойством, которое мы приписываем электронной комплементарности. Однако такая пара, как A — U, может образовываться не единственным способом (имидазольная структура, обратная имидазольная структура или уотсон-криковская структура), и формирование двойной уотсон-криковской спирали должно обеспечиваться действием дополнительных геометрических факторов и среды. (О некоторых других взаимодействиях при спаривании оснований см. дополнение 6.1.)

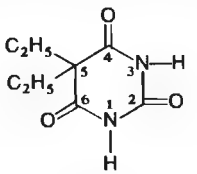
6.7. Стэкинг оснований

Анализ кристаллической структуры спиральных молекул нуклеиновых кислот показывает, что плоскости следующих друг за другом оснований практически параллельны и в значительной мере перекрываются; двугранный угол между соседними плоскостями не превышает 20° (эти особенности отчетливо видны в структуре двойной спирали ДНК на рис. 3.12). Является ли такое стопочное расположение (стэкинг) оснований следствием геометрических ограничений, налагаемых на углы вращения остова специфической гео-

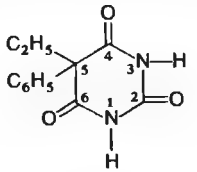
Дополнение 6.1

НЕКОТОРЫЕ ДРУГИЕ СЛУЧАИ СПАРИВАНИЯ ОСНОВАНИЙ

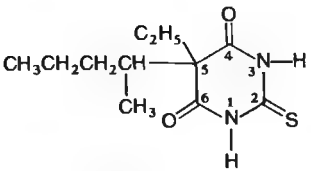
Весьма интересно в данном контексте рассмотреть некоторые свойства барбитуратов — соединений, близких по своему строению к урацилу. Эти лекарственные вещества широко применяются как снотворные, уменьшая потребление кислорода в тканях млекопитающих и подавляя процесс дыхания в митохондриях. Ниже приведены структурные формулы барбитала, фонобарбитала, тиопентала и урацила.



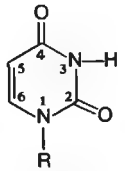
Барбитал
($1\,000\ M^{-1}$)



Фенобарбитал
($1\,200\ M^{-1}$)



Тиопентал
($600\ M^{-1}$)



Урацил
($100\ M^{-1}$)

Главное различие между барбитуратами касается боковой цепи, связаиной с C^5-атомом, и наличия кислорода или серы у атома C^2. Из-за сходства этих соединений с урацилом можно ожидать, что барбитураты будут образовывать специфические комплексы с аденином, что и происходит в действительности. Под каждой структурной формулой в скобках указаны константы их ассоциации с аденином в дейтерохлороформе (Kyogoku et al., 1968). С цитозином, гуанином и урацилом эти соединения не взаимодействуют, а константы их ассоциации между собой лежат в интервале от 1 до $10\ M^{-1}$.

Взаимодействие между аденином и барбитуратами значительно сильнее, чем между аденином и урацилом. Это обусловлено двумя факторами. Во-первых, NH-группа барбитуратов является намного более кислой: для уридина $pK = 9{,}4$, в то время как для барбитала $pK = 7{,}8$, для фенобарбитала 7,3, а для тиопентала 7,4. Во-вторых, барбитураты несут вдвое большее число NH- и C=O-групп, чем урацил, так что чисто статистический эффект увеличивает константу связывания барбитуратов с аденином в два раза. Заметим, что комплекс аденина с тиопенталом оказывается менее прочным, чем с фенобарбиталом и барбиталом, предположительно из-за меньшей электроотрицательности серы по сравнению с кислородом.

Полагают, что барбитураты становятся активными в неполярном окружении, возможно в мембранах. Поэтому результаты, полученные в хлороформе, могут не слишком отличаться от таковых для физиологических условий, в которых эти вещества взаимодействуют с основаниями нуклеотидов. Конечно, вопрос о том, имеют ли такого рода данные отношение к механизму фармакологического действия барбитуратов, остается открытым, но приведенные результаты показывают, как физические исследования механизма химических процессов могут перерасти в изучение механизма физиологического явления.

метрией уотсон-криковского спаривания оснований, или это расположение энергетически более выгодно и само по себе служит стабилизирующим фактором для уотсон-криковской структуры?

СТЭКИНГ ОСНОВАНИЙ В ОДНОЦЕПОЧЕЧНОЙ poly (A) В РАСТВОРЕ

Условия, необходимые для стэкинга оснований, интенсивно исследовались путем анализа свойств одноцепочечных полинуклеотидов в растворе, особенно полиаденциловой кислоты. Наиболее плодотворным оказалось в этом отношении изучение оптических свойств растворов. На рис. 6.11 сопоставлены спектры поглощения AMP и poly (A) в растворах, содержащих эквивалентное количество нуклеотидных остатков. Интегральное поглощение poly (A) существенно меньше, чем поглощение составляющих ее мономеров, — эффект, известный под названием гипохромизма (см. гл. 7, где обсуждаются оптические явления). Кроме того, длина волны, отвечающая максимуму поглощения poly (A) при нейтральных pH, на 3 нм меньше, чем соответствующая величина для AMP. Однако эта значительная разница в поглощении постепенно уменьшается при повышении температуры.

На рис. 6.12 приведены данные, полученные М. Ленгом и Г. Фельзенфельдом (1966 г.). Представлена зависимость поглощения при 260 нм от температуры для poly (A) и для различных олигомеров этого основания. Оптическая плотность ограничена двумя предельными значениями — оптической плотностью для гипохромной формы, существующей при очень низких температурах (~ −20°C), и соответствующим значением для высокотемпературной формы (~ 90°C) с оптической плотностью, характерной для AMP. В том же температурном интервале поглощение AMP почти не меняется.

Изучение других спектроскопических свойств (таких, как круговой дихроизм) дает аналогичные результаты и свидетельствует о том, что в poly (A) происходят температурные изменения, не наблюдающиеся в AMP. Объяснить все эти данные можно тем, что стэкинг оснований постепенно уменьшается с ростом температуры (см. гл. 7, где обсуждаются физические основы гипохромизма в стопочных структурах).

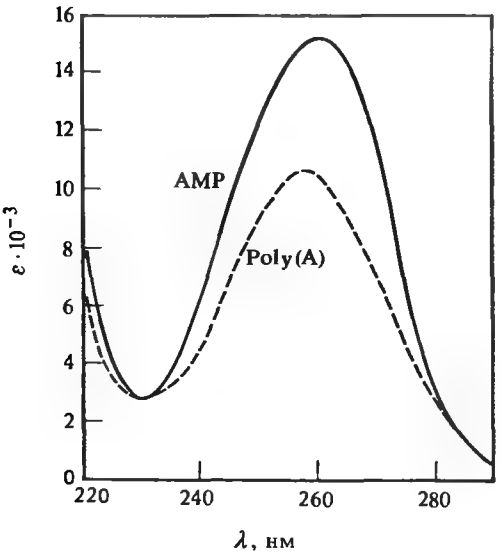


РИС. 6.11. Спектры поглощения AMP и poly (A); отчетливо выражен гипохромизм. [K.E. Van Holde, Physical Biochemistry (Englewood Cliffs, N.J.: Prentice-Hall, 1971), p. 168.]

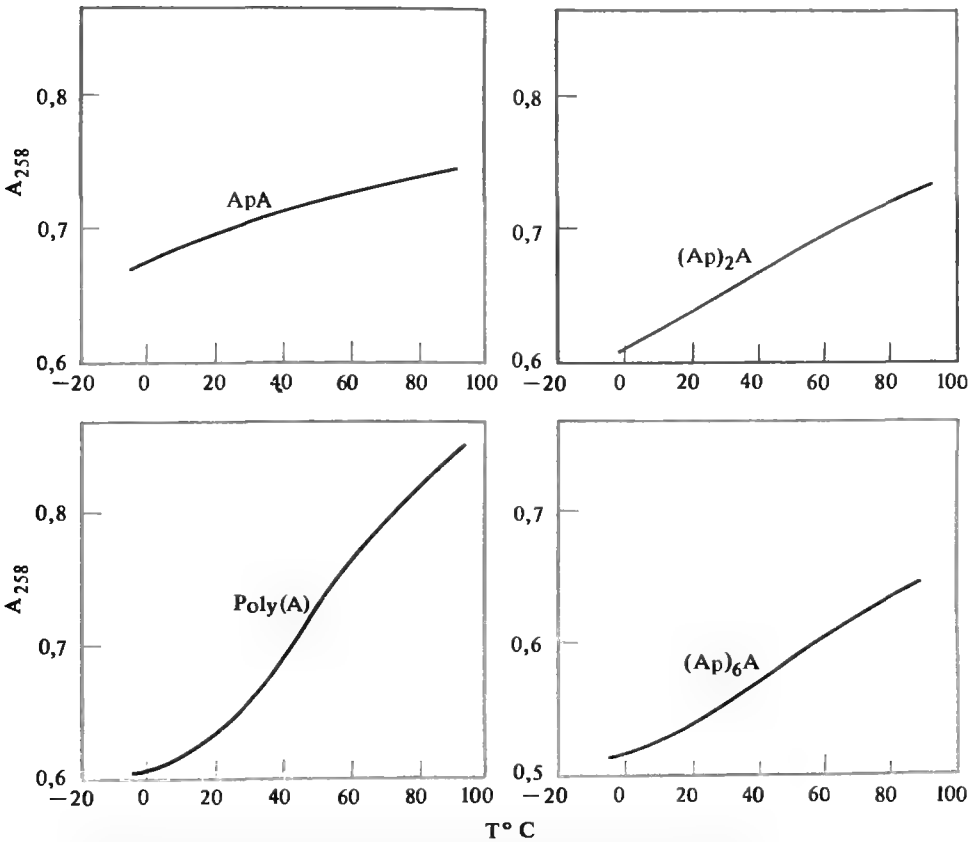


РИС. 6.12. Зависимость поглощения при 258 нм от температуры для poly (A) и различных олигомеров. [M. Leng, G. Felsenfeld, J.Mol.Biol.,15, 455 1966).]

Данные рис. 6.12 поддаются прямой интерпретации. Предположим, что существует равновесие между двумя состояниями динуклеотида, и будем рассматривать цепь из N нуклеотидных звеньев как набор из $N - 1$ динуклеотидной единицы, стэкинг в каждой из которых не зависит от состояния других единиц. Обозначим через K константу равновесия для рассматриваемого процесса, а через f_u — долю единиц, в которых отсутствует стэкинг. Тогда

$$K = f_u/(1 - f_u) \tag{6.1а}$$

и

$$f_u = (A_T - A_{min})/(A_{max} - A_{min}) \tag{6.1б}$$

где A_{max} — высокотемпературный предел поглощения, A_{min} — низкотемпературный предел, A_T — поглощение при температуре T. Для определения термодинамических характеристик полученную таким образом зависимость K от температуры нужно построить в координатах $\ln K$ от $1/T$ (рис. 6.13), в которых график представляет собой прямую. Из приведенных данных получаем $\Delta H° = 13$ ккал на моль (стэкингов) и $\Delta S° = 40$ э.е. на моль для процесса разрушения стэкинга. (Следует отметить, однако, что недавно в результате калориметрического анализа были получены существенно отличающиеся зна-

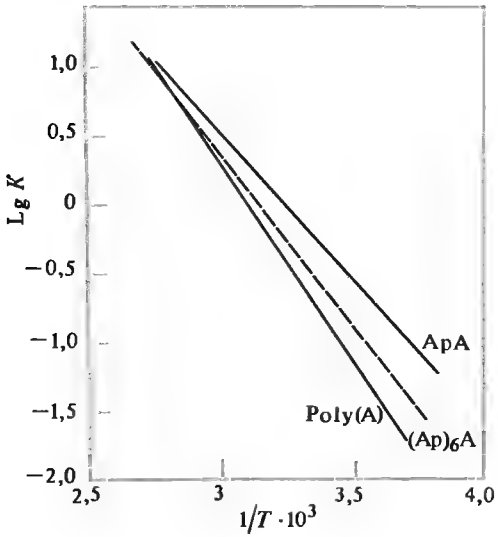


РИС. 6.13. Зависимость lg K от $1/T$ для олигомеров и полимеров аденина. [M. Leng, G. Felsenfeld, J.Mol.Biol., 15, 455 (1966).]

чения термодинамических параметров; см. Дополнение 6.2.) Такие же в первом приближении величины были найдены и для более коротких цепей, что подтверждает гипотезу о почти независимом образовании каждого стэкинга.

Дополнение 6.2

РАСХОЖДЕНИЕ ДАННЫХ ПО ТЕРМОДИНАМИКЕ РАЗРУШЕНИЯ СТЭКИНГА

К. Бреслауэр и Дж. Стюртевант (1977 г.) исследовали гептамер аденилата методом дифференциальной сканирующей калориметрии и получили $\Delta H = 3{,}4$ ккал для разрушения одного моля стэкингов. Они также отметили, что вантгоффова эктальпия, вычисленная из кривых плавления, полученных оптическими методами, меняется в интервале от 5 до 13 ккал · моль$^{-1}$. Расхождения могут быть обусловлены трудностями в получении аккуратной нулевой линии при оптических измерениях. Ясно, что этот вопрос требует дальнейших исследований (см. Дополнение 21.1 о калориметрии).

СТЭКИНГ МОНОНУКЛЕОЗИДОВ В РАСТВОРЕ

Склонность оснований к образованию стопок проявляется не только в наличии взаимодействия между ближайшими соседями в полинуклеотидной цепи. Имеются убедительные данные о том, что мононуклеозиды тоже образуют агрегаты в водных растворах. Как инфракрасные спектры, так и спектры ЯМР указывают, что эта агрегация не связана с образованием водородных связей. Более того, сигналы ЯМР, соответствующие индивидуальным протонам в комплексе, всегда сдвинуты в высокочастотную область, что и должно быть при влиянии кольцевых токов на положение этих сигналов (см. гл. 9 о ЯМР). Таким образом, довольно сильная тенденция мононуклеозидов к образованию стопок в полиадениловой кислоте является свойством самих оснований.

Одним из наиболее плодотворных методов исследования термодинамических свойств процесса образования стопок из мономеров является классический метод измерения давления паров над разбавленным раствором мононуклеозидов. Считая, что раствор при низких концентрациях мононуклеозидов близок к идеальному, измеряют понижение давления паров растворителя, вызываемое введением в раствор оснований. Относительное понижение давления равно, согласно закону Рауля, мольной доле всех растворенных веществ:

$$(P_0 - P_\nu)/P_0 = \chi \tag{6.2}$$

где P_0 и P_ν — давление паров растворителя в отсутствие и при наличии растворенного вещества соответственно, χ — мольная доля всех растворенных веществ в растворе. Таким образом, если n_0 — число молей растворителя, n_1 — число молей растворенных мономеров, n_2 — число молей растворенных димеров (пар, связанных межплоскостным взаимодействием), n_3 — число молей растворенных тримеров (трех уложенных в стопку мономерных единиц) и т.п., то

$$\chi = \left(\sum_{i \geqslant 1} n_i\right) \Big/ \sum_{i \geqslant 0} n_i \cong \left(\sum_{i \geqslant 1} n_i\right) \Big/ n_0 \tag{6.3}$$

поскольку для разбавленных растворов $n_0 \gg \sum_{i \geqslant 1} n_i$. Зная мольную долю χ, можно теперь вычислить суммарную моляльность всех растворенных соединений

$$m = \sum_{i \geqslant 1} m_i \tag{6.4а}$$

где m_i — моляльность i-го компонента.

Величину m имеет смысл сравнить с известной суммарной моляльностью m' всех мономерных единиц в растворе безотносительно к степени их ассоциации:

$$m' = \sum_{i \geqslant 1} i m_i \tag{6.4б}$$

Отношение m/m' известно как эффективный (practical) осмотический коэффициент ϕ_p. Ясно, что $\phi_p \leqslant 1$ и тем меньше, чем больше степень ассоциации мономерных единиц (см. гл. 25, где обсуждаются вопросы осмотического давления).

Не составляет труда связать эффективный осмотический коэффициент с соответствующими параметрами модели процесса ассоциации мономерных единиц B так, как это было сделано Т. Соли и Дж. Шеллманом (1968 г.). Для неограниченной ассоциации мономеров B_1 мы имеем

$$\begin{gathered} 2B_1 \rightleftarrows B_2, K_2 \\ B_1 + B_2 \rightleftarrows B_3, K_3 \\ B_1 + B_3 \rightleftarrows B_4, K_4 \end{gathered} \tag{6.5}$$

где K_i — константа ассоциации для B_i. Для случая, когда происходит лишь димеризация, легко получить соотношение

$$(1 - \phi_p)/(2\phi_p - 1)^2 = K_2 m' \tag{6.6}$$

Аналогичным образом для случая, когда происходит бесконечная ассоциация, но все K_i одинаковы (все $K_i = K$), имеем

$$(1 - \phi_p)/\phi_p^2 = Km' \tag{6.7}$$

(см. задачу 6.1). Уравнения (6.6) и (6.7) позволяют непосредственно использовать эффективные осмотические коэффициенты в моделях ассоциации. На рис. 6.14,*А* приведен график зависимости $(1 - \phi_p)/(2\phi_p - 1)^2$ от m' для дезоксиаденозина. Из этих данных видно, что при $m' = 0{,}03$ зависимость начинает заметно отклоняться от линейной, что указывает

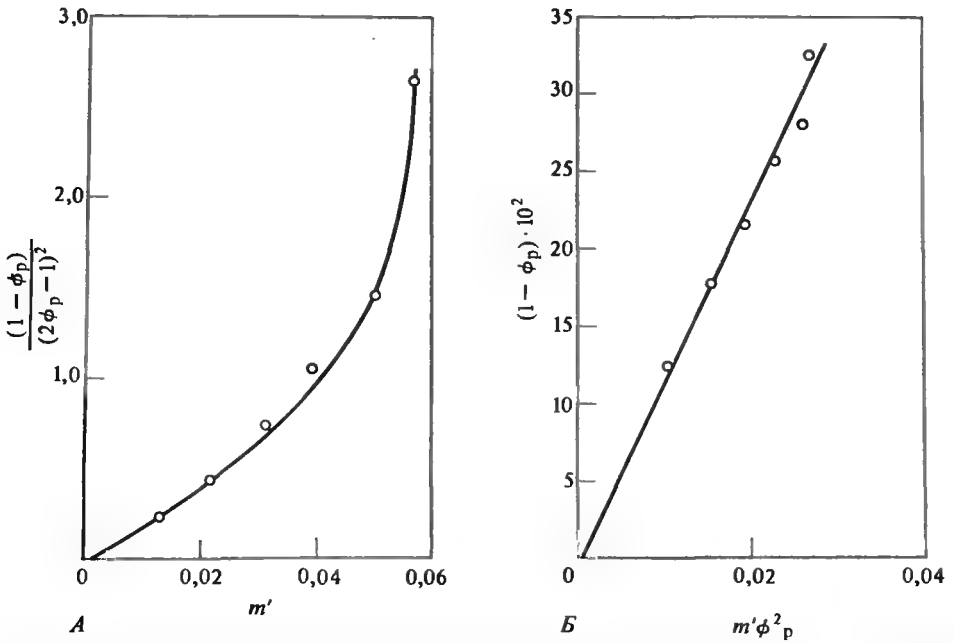


РИС. 6.14. Осмотическое давление над раствором дезоксиаденозина. *А*. Зависимость $(1-\phi_p)/(2\phi_p-1)^2$ от m'. *Б*. Зависимость $(1-\phi_p)$ от $m'\phi_p^2$. [T.N. Solie, J.A. Schellman, J.Mol.Biol., **33**, 61 (1968).]

на присутствие агрегатов более высокого порядка. Предельный наклон при $m' \to 0$ дает константу ассоциации для димеризации. Если перестроить данные в координатах $\{m'\phi_p^2;\ 1-\phi_p\}$ [уравнение (6.7)], то экспериментальные точки лягут на прямую (рис. 6.14,*Б*). Это позволяет сделать вывод, что образование всех агрегатов более высокого порядка характеризуется близкими константами ассоциации.

ПУРИНОВЫЙ СТЭКИНГ СИЛЬНЕЕ, ЧЕМ ПИРИМИДИНОВЫЙ

В табл. 6.4 приведены некоторые константы ассоциации, полученные описанными методами. Вообще говоря, все константы ассоциации малы и указывают на то, что сила взаимодействия пропорциональна размеру колец, которые участвуют в стэкинге. Константы ассоциации образуют ряд $K_{\text{пурин — пурин}} > K_{\text{пурин — пиримидин}} > K_{\text{пиримидин — пиримидин}}$.

Свободную энергию, значения которой приведены в табл. 6.4, нельзя прямо сопоставлять с энергией стэкинга оснований в полинуклеотидной цепи, поскольку первая из них вычислена с учетом добавочного члена $-RT\ln 55{,}6$ (см. гл. 5). Таким образом, для более корректного сопоставления ко всем величинам, приведенным в таблице, следует прибавить $-RT\ln 55{,}6 = -2{,}4$ ккал · моль$^{-1}$. Так, для дезоксиаденозина получаем при 25 °С величину $-3{,}9$ ккал · моль$^{-1}$. $\Delta G°$ стэкинга аденина в poly(A) составляет -1 ккал · моль$^{-1}$ при той же температуре и несколько отличных прочих условиях. Можете ли вы высказать какие-либо соображения о причинах различия этих величин?

Таблица 6.4

КОНСТАНТЫ АССОЦИАЦИИ И ЗНАЧЕНИЯ СВОБОДНОЙ ЭНЕРГИИ ПРЕДПОЛАГАЕМОГО СТЭКИНГА ОСНОВАНИЙ В ВОДНОМ РАСТВОРЕ ПРИ 25°C[1)]

Осиование	K, Мл$^{-1}$	ΔG, ккал · моль$^{-1}$	Основание	K, Мл$^{-1}$	ΔG, ккал · моль$^{-1}$
dA	12	−1,50	dA — dT	3 — 6	От −0,70 до −1,10
Пуринрибозид	3,5	−0,74	dG — dC	4 — 8	От −0,80 до −1,20
I	3,0	−0,64	dA — dC	3 — 6	От −0,70 до −1,10
dT	0,91	+0,06	dA — dU	3	−0,70
dC	0,91	+0,06	dT — dC	0,91	+0,06
U	0,70	+0,21			

[1)] T. N. Solie, J. A. Schellman, J. Mol. Biol., 33, 61 (1968).

РОЛЬ РАСТВОРИТЕЛЯ В СТАБИЛИЗАЦИИ СТЭКИНГА

Какова природа сил, обеспечивающих стэкинг-взаимодействие? Во-первых, совершенно очевидно, что отчасти стэкинг обусловлен взаимодействием между индуцированнными диполями, образованными π-электронами оснований. Далее, важный факт, проливающий свет на природу стабилизации стопочных структур, заключается в том, что эти структуры наиболее стабильны в водных растворах. К этому выводу приводят, например, результаты исследования изменения гипохромизма ДНК в различных растворителях, но, по-видимому, более прямое подтверждение можно получить при изучении фотохимического образования тиминовых димеров в динуклеотидах тимидил—тимидин (ТрТ). Хорошо известно, что облучение ДНК УФ-светом с длиной волны 280 нм приводит к ковалентному сшиванию соседних тиминовых остатков в одной цепи. Накопленные данные позволяют сделать вывод, что для осуществления димеризации основания должны примыкать плоскостями друг к другу. Было обнаружено, что при данной дозе облучения количество образованных из ТрТ димеров очень сильно зависит от растворителя: больше всего димеров T_D образуется в воде. Количество димеров, образующихся в неводных растворителях, оказывается пропорциональным вероятности образования межплоскостного контакта в ТрТ для этих сред. В табл. 6.5 суммированы данные для некоторых растворителей, перечисленных в порядке увеличения их способности содействовать межплоскостному взаимодействию. Видно, что этот порядок в точности соответствует той последовательности, в которой увеличивается количество димеров T_D, образующихся в этих растворителях.

Приведенные результаты указывают на важную роль водной среды в стабилизации упорядоченных конформаций нуклеиновых кислот. Здесь можно провести аналогию с белками, где гидрофобные взаимодействия тоже занимают центральное место в формировании упорядоченных структур. Таким образом, весьма незначительные на первый взгляд молекулы воды играют очень большую биологическую роль.

Итак, вода играет важную роль в стабилизации стэкинга. Но мы не можем безоговорочно утверждать, что гидрофобные взаимодействия — это главные силы, стабилизирующие стопкообразные структуры неполярных молекул в воде. Эти взаимодействия являются эндотермическими и, следовательно, определяются энтропией. В то же время стэкинг-взаимодействие оказывается экзотермическим. Чтобы глубже понять природу сил, стабилизирующих стэкинг, Д. Крозерс и Д. Ратнер (1968 г.) исследовали термодинамику межплоскостного взаимодействия дезоксигуанозина и актиномицина в растворе. Структура актиномицина показана на рис. 23.24; в этой молекуле имеется феноксазоновое

Таблица 6.5

ОБРАЗОВАНИЕ ТИМИНОВЫХ ДИМЕРОВ В РАЗЛИЧНЫХ РАСТВОРИТЕЛЯХ [1,2]

Растворитель	Количество образовавшихся димеров T_D, %	Растворитель	Количество образовавшихся димеров T_D, %
трет-Бутанол	5	Этиленгликоль	19
Этанол	5,5	Формамид	21
Метанол	6	Глицерин	33
н-Бутанол	7,5	Вода	35

[1] W. Wacker, Angew Chem. Internat., **4**, 150 (1965).

[2] Растворители перечислены в порядке увеличения их способности содействовать межплоскостному взаимодействию в ТрТ.

кольцо, которое, как показывают кристаллографические данные Г. Собелла, может вступать в межплоскостное взаимодействие с кольцом гуанина. Крозерс и Ратнер установили, что в водном фосфатном буфере при 25°C образование комплекса характеризуется изменением энтальпии $\Delta H^0 = -10{,}3$ ккал · моль$^{-1}$ и изменением энтропии $\Delta S^0 = -19{,}8$ кал · моль$^{-1}$ · град$^{-1}$, но при добавлении метанола и ΔH^0, и ΔS^0 увеличиваются по абсолютной величине. Например, при концентрации метанола 25% $\Delta H^0 = -14{,}05$ ккал · моль$^{-1}$, а $\Delta S^0 = -33{,}7$ кал · моль$^{-1}$ · град$^{-1}$.

Таким образом, в чисто водном растворе имеются скрытая положительная энтропия и положительная энтальпия образования комплекса, которые теряются при добавлении метанола. Весьма вероятно, что этот скрытый вклад в энтальпию и энтропию возникает за счет гидрофобных эффектов, включающих перегруппировку молекул воды (см. гл. 5). Другой подход к анализу межплоскостных взаимодействий дан в дополнении 6.3.

Дополнение 6.3.

ВЛИЯНИЕ ПОВЕРХНОСТНОГО НАТЯЖЕНИЯ

Интересный подход к исследованию энергетики стэкинга оснований разработали Синаноглу и Абдулнур (1964 г.). Они развили идею о том, что поверхностное натяжение наряду с другими факторами может быть использовано для вычисления работы, необходимой для образования из сольватированных неупорядоченных оснований спиральной структуры, что сопровождается вытеснением растворителя из промежутка между основаниями. Задача в этом случае сводится к вычислению работы, которая необходима для создания полостей в растворе для индивидуальных сольватированных оснований и для оснований, уложенных в спираль. Работа, затраченная на создание полости площадью dA, есть просто $dw = \gamma dA$, где γ — поверхностное натяжение. Очевидно, для создания отдельных полостей для каждого основания необходимо совершить больше работы, чем для создания одной большой полости, которая вместит все основания без растворителя между ними. Следовательно, будет наблюдаться тенденция (до тех пор, пока рассматривается только взаимодействие растворитель — растворитель) к упаковке оснований в одну полость.

Но почему вода проявляет такую относительно сильную тенденцию к стабилизации стопочных структур? Возможно, дело в том, что она обладает относительно высоким поверхностным натяжением — около 72 дин · см$^{-2}$ — по сравнению, например, с этанолом, для которого эта величина составляет около 22 дин · см$^{-2}$ при комнатной температуре. Разница в энергии, необходимой для создания отдельных полостей и образования одной большой полости, для воды оказывается

больше, чем для других растворителей. Растворители, вызывающие денатурацию ДНК (метанол, этиленгликоль, формамид, глицерин, энатнол, *н*-пропанол и *н*-бутанол), значительно уступают воде в способности содействовать образованию спирали за счет эффекта поверхностного натяжения. Эти растворители, однако, имеют также более низкую диэлектрическую проницаемость и поэтому увеличивают электростатическую дестабилизацию спиральных молекул нуклеиновых кислот.

6.8. Третичная структура нуклеиновых кислот

Все рассмотренные нами данные дают довольно полное представление о геометрических факторах и силах, которые определяют вторичную структуру нуклеиновых кислот. На основании этих данных можно прийти к выводу, что нуклеиновые кислоты находятся в основном в двух конформационных состояниях — в виде беспорядочного клубка и линейных одно- и двухцепочечных спиралей. Такое впечатление ошибочно и вызвано отчасти чрезмерным увлечением модельными системами и убеждением, что третичные взаимодействия играют важную роль лишь в белках.

Хорошо известным примером нуклеиновой кислоты, имеющей четко выраженную третичную структуру, является транспортная РНК. Эти одноцепочечные полимерные молекулы состоят обычно из 73 — 93 нуклеотидов и складываются таким образом, что их вторичная структура имеет форму клеверного листа (рис. 3.6). тРНК образуют семейство молекул, каждая из которых специфична к определенной аминокислоте. Эта специфичность связана со способностью 3′-конца каждой цепи тРНК образовывать эфирную связь с одной определенной аминокислотой с помощью специфического фермента. Аминокислота, которая может быть присоединена к данной тРНК, соответствует определенному триплету в последовательности оснований тРНК — антикодону, комплементарному кодону. Место присоединения данной аминокислоты к растущей полипептидной цепи определяется взаимодействием между антикодоном тРНК и кодоном матричной РНК.

Согласно многочисленным данным, вторичная структура тРНК имеет форму клеверного листа, состоящего из спиральных участков и расположенных между ними петель. Образование такой структуры определяется в значительной мере сильной тенденцией спаренных и уложенных в стопки оснований к образованию спиралей. В то же время имеется много данных, полученных для растворов, которые ясно указывают на наличие у тРНК четко выраженной третичной структуры. Природу этой укладки удалось установить в основном благодаря кристаллографическим исследованиям трехмерной структуры дрожжевой фенилаланиновой тРНК, выполненным Ричем, Кимом, Клугом и их коллегами.

На рис. 3.17 представлена схема укладки «клеверного листа» в третичную структуру. Молекула имеет Г-образную форму; один из «рукавов» буквы Г образован спиралью акцепторного стебля, непрерывно переходящей в спираль псевдоуридилового стебля, а другой — антикодоновым и дигидроуридиловым стеблями. Угол буквы Г образован ТΨС- и дигидроуридиловой петлями. На концах буквы Г находятся антикодоновая петля и 3′-акцепторный конец, разделенные расстоянием около 75 Å. Из этого грубого описания ясно, что антикодоновая петля находится «снаружи» структуры и поэтому может участвовать в спаривании оснований. Хотя существование этой конформации в растворе еще не доказано, большинство данных согласуется с ней.

Третичная структура тРНК частично стабилизирована ранее неизвестными и потому представляющими большой интерес взаимодействиями, включающими образование водородных связей в комплексах из трех оснований (рис. 3.18). В результате этих необычных взаимодействий образуются мостики между различными участками цепи и стабилизируется более компактная конфигурация. Ясно также, что в стабилизации структуры участвуют поливалентные катионы (Mg^{2+}, полиамины), хотя детали этих взаимодействий до конца не выяснены.

Результаты, полученные при изучении транспортных РНК, показывают, что нуклеиновые кислоты могут иметь высокоупорядоченную трехмерную конформацию. Не вызывает сомнения, что гораздо более крупные молекулы РНК, такие, как рибосомные РНК, также имеют сложную трехмерную структуру. Существуют и молекулы ДНК сложной формы. О механизме образования третичной структуры нуклеиновых кислот известно еще слишком мало, и проблема предсказания уникальной трехмерной структуры этих молекул исходя из их нуклеотидной последовательности представляется невероятно трудной (если вообще разрешимой), так же как и аналогичная проблема для белков. (Дальнейшее обсуждение третичной структуры нуклеиновых кислот можно найти в гл. 24.).Тем не менее ясно, что основными факторами, благодаря которым формируется третичная структура полинуклеотидов и белков, являются геометрические и стерические ограничения, водородные связи, гидрофобные взаимодействия и электрические силы.

Краткие выводы

Конформация остова полинуклеотидной цепи определяется шестью углами внутреннего вращения, имеющимися в каждой повторяющейся единице цепи. Наличие большого числа переменных усложняет стереохимический анализ конформаций цепи, но путем последовательных шагов можно значительно упростить задачу и показать, что стерические эффекты существенно уменьшают количество допустимых конформаций.

Два основных типа взаимодействий, стабилизирующих структуру нуклеиновых кислот, связаны с образованием пар и стэкингом оснований. Возможно существование разных типов пар оснований, и некоторые из них действительно наблюдались при кристаллографическом исследовании мономеров. Для кристаллов мономерных производных аденина и урацила образование уотсон-криковских пар (как в двойной спирали ДНК) никогда не наблюдалось, но часто удавалось обнаружить другие типы пар. Для комплементарных же динуклеотидов ApU, кристаллическая структура которых определена с атомным разрешением, характерно образование уотсон-криковской спирали. Спаривание оснований было исследовано и для растворов мономеров; установлено, что A спаривается в основном с U, а G с C, хотя при образовании специфического комплекса A — U реализуется более одной схемы спаривания. Вся совокупность данных свидетельствут о том, что уотсон-криковская спираль образуется вследствие «электронной» комплементарности в A—U- и G—C-парах и вследствие геометрических ограничений в двойной спирали.

Стэкинг-взаимодействие проявляется в водных растворах и наблюдается в одноцепочечных полинуклеотидах, таких, как poly (A), и между свободными мономерами нуклеозидов. Вообще говоря, пурины охотнее участвуют в стэкинге, чем пиримидины. Взаимодейтвие связано с отрицательной величиной ΔH, что отличает его от гидрофобных взаимодействий, хотя другие данные указывают на значительную роль в стабилизации стэкинга оснований гидрофобного эффекта.

Нуклеиновые кислоты могут иметь также и специфическую третичную структуру, как видно на примере транспортной РНК. Компактная специфическая структура тРНК формируется благодаря образованию необычных пар оснований и наличию других взаимодействий.

ЗАДАЧИ

6.1. Используя определение эффективного осмотического коэффициента ϕ_p и уравнения (6.4) и (6.5), докажите справедливость уравнений (6.6) и (6.7).

6.2. Экспериментатор измеряет давление паров адеиозина (A) над тремя разными растворами этого соединения. Первый растворитель — это 8 М водный раствор мочевины, вто-

рым растворителем является CCl_4, а третий представляет собой чистую воду. Измерения проводятся для того, чтобы установить активность аденозина (или концентрацию свободного аденозина) в этих трех растворах. Затем экспериментатор сопоставляет результаты измерений с общим количеством аденозина $(A)_0$, добавленным в раствор. Он проводит эти опыты в широком диапазоне концентраций А и приходит к выводу, что для раствора I соотношение между $(A)_0$ и действительной концентрацией свободного аденозина (A) в растворе соответствует равенству $(A)_0 = (A)$, т.е. в растворе I концентрация свободных мономеров равна суммарной концентрации. Для раствора II оказывается, что $(A)_0 = (A) + C(A)^2$, где C — константа, причем согласие с результатами измерений не улучшается добавлением к правой части соотношения членов, соответствующих более высоким степеням (A), чем $(A)^2$. Для раствора III

$$(A)_0 = C_1(A) + C_2(A)^2 + C_3(A)^3 + C_4(A)^4 + \ldots$$

где $C_1 = 1$, C_2 — константа и

$$C_k = (k/2^{k-1})C_2^{k-1} \quad \text{и} \quad k = 3, 4, \ldots$$

Исходя из этих данных опишите типы специфических взаимодействий, которые реализуются в указанных растворах. Для обоснования ваших предположений покажите, что приведенные здесь уравнения однозначно следуют из этих предположений.

6.3. В процессе изучения стэкинга оснований мононуклеозидов экспериментатор измерил параметр ϕ_p при трех значениях суммарной концентрации мононуклеозидов и обнаружил, что величина K (константа равновесия для стэкинга), вычисленная из этих данных, почти не зависит от того, какую модель он использует — ту, в которой предполагается образование лишь димеров, или же модель неограниченной ассоциации, в которой предполагается образование агрегатов высших порядков (для всех агрегатов характерна одна и та же константа равновесия K). Экспериментатор считает, что где-то допущена ошибка, но не может найти ее. Если вы считаете, что ошибка действительно допущена, укажите, в чем несовместимость полученных данных. В противном случае укажите условия, при которых вы могли бы получить почти одинаковые значения K из данных, соответствующих трем разным концентрациям. (*Примечание*: тривиальный случай $K = 0$ не рассматривается.)

6.4. Для двухцепочечных нуклеиновых кислот наблюдается переход спираль—клубок, который до некоторой степени аналогичен плавлению poly(A) с ростом температуры. Кривая плавления ДНК является, однако, чрезвычайно крутой. Середина этого кооперативного перехода характеризуется температурой плавления T_m. Основываясь на данных этой главы, укажите, какие факторы оказывают существенное влияние на крутизну кривой и температуру плавления ДНК. Почему? Можете ли вы дать физическое объяснение кооперативности перехода?

6.5. Рассмотрим такой реагент, как мочевина. Известно, что в ее присутствии происходит эффективная денатурация белков. Какое влияние, по вашему мнению, окажут водные растворы мочевины с концентрацией от 1 до 8 М на стабильность упорядоченных одно- и двухцепочечных нуклеиновых кислот? Почему?

ЛИТЕРАТУРА

Общая

Olson W. K., Flory P. J. 1972. Steric configurations of polynucleotide chains, I: Steric interactions in polynucleotides: A virtual bond model, Biopolymers, **11**, 1 [В этой и других статьях того же выпуска журнала приведен критический анализ данных о конформации и стереохимии полинуклеотидов.]

Ts'o P. O. P. ed., 1974. Basic Principles in Nucleic Acid Chemistry, 2 vols., New York, Academic Press. [Изложены основные вопросы физики и химии нуклеиновых кислот; авторы — крупные специалисты в данной области. В гл. 6 1-го тома рассмотрены свойства мономерных единиц и их взаимодействие.]

Voet D., Rich A., 1970. The crystal structure of purines, pyrimidines and their intermolecular complexes, Progr. Nucleic Acid Res. Mol. Biol., **10**, 183. [Содержит прекрасную сводку данных о комплексах между основаниями, образующихся с помощью водородных связей, в кристаллическом состоянии.]

Специальная

Breslauer K. J., Sturtevant J. M., 1977. Calorimetric investigation of base-stacking in Ribo A_7, Biophys. Chem., **7**, 205.

Crothers D. M., Ratner D. I., 1968. Thermodynamic studies of a model system for hydrophobic bonding, Biochemistry, **7**, 1823.

Hoogsteen K., 1959. The structure of crystals containing a hydrogen-bonded complex of 1-methylthymine and 9-methyladenine, Acta Cryst., **12**, 822.

Hoogsteen K., 1963. The crystal and molecular structure of a hydrogen-bonded complex between 1-methylthymine and 9-methyladenine, Acta Cryst., **16**, 907.

Kim S. H., Quigley G. J., Suddath F. L., McPherson A., Sneden D., Kim J. J., Weinzierl J., Rich A., 1973. Three-dimensional structure of yeast phenylalanine transfer RNA: Folding of the polynucleotide chain, Science, **179**, 285.

Kim S. H., Suddath F. L., Quigley G, J., McPherson A., Sussman J. L., Wang A. H. J., Seeman N. C., Rich A., 1974. Three-dimensional tertiary structure of yeast phenylalanine transfer RNA, Science, **185**, 435.

Kyogoku Y., Lord R. C., Rich A., 1968. Specific hydrogen bonding of barbiturates to adenine derivatives, **218**, 69.

Lend M., Felsenfeld G., 1966. A study of polyadenylic acid at neutral pH, J. Mol. Biol., **15**, 455.

Pullman B., Claverie P., Caillet J., 1966. Van der Waals—London interactions and the configuration of hydrogen-bonded purine and pyrimidine pairs, Proc. Natl. Acad. Sci. USA, **55**, 904.

Rich A., Kim S.H., 1978. The three-dimensional structure of transfer RNA Sci. American, **238** (1), 52.

Robertus J. D., Ladner J. E., Finch J. T., Rhodes D., Brown R. S., Clark B. F. C., Klug A., 1974. Structure of yeast phenylalanine tRNA at 3 Å resolution, Nature, **250**, 546.

Rosenberg J. M., Seeman N. C., Kim J. J. P., Suddath F. L., Nicholas H. B., Rich A., 1973. Double helix at atomic resolution, Nature, **243**, 150.

Sinanoğlu O., Abdulnur S., 1964. Hydrophobic stacking of bases and the solvent denaturation of DNA, Photochem. Photobiol., **3**, 333.

Solie T. N., Schellman J. A., 1968. The interaction of nucleosides in aqueous solution, J. Mol. Biol., **33**, 61.

Приложение А

Основы матричной алгебры

Матрицей называется таблица, элементами которой являются числа или символы. Например,

$$\begin{pmatrix} 8 & 7 \\ 23 & 28 \end{pmatrix}, \quad \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

— это матрицы размером 2 × 2 (содержащие по две строки и по два столбца); символ a_{12} обозначает элемент, принадлежащий первой строке и второму столбцу. В общем случае a_{ij} представляет собой элемент i-й строки и j-го столбца. Для левой матрицы $a_{22} = 28$, $a_{21} = 23$ и т.д. В этой книге матрицы обозначаются буквами, выделенными жирным шрифтом со значком «тильда» под ними. Например, мы можем представить правую из приведенных выше матриц как $\underset{\sim}{\mathbf{A}}$, где

$$\underset{\sim}{\mathbf{A}} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

Матрица-строка (называемая также вектор-строкой) состоит только из одной строки:

$$\underset{\sim}{\mathbf{A}}_{\mathrm{r}} = (a_{11}, a_{12})$$

$\underset{\sim}{\mathbf{A}}_{\mathrm{r}}$ — это матрица размером 1 × 2 (с одной строкой и двумя столбцами). *Матрица-столбец* (или вектор-столбец) состоит только из одного столбца:

$$\underset{\sim}{\mathbf{A}}_{\mathrm{c}} = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}$$

$\underset{\sim}{\mathbf{A}}_{\mathrm{c}}$ — это матрица размером 2 × 1 (с двумя строками и одним столбцом). В общем случае матрица может быть любого размера — с n строками и m столбцами.

Перемножение матриц

Две матрицы могут быть перемножены, так что получится третья матрица:

$$\underset{\sim}{\mathbf{a}}\,\underset{\sim}{\mathbf{b}} = \underset{\sim}{\mathbf{c}}$$

Элемент c_{ij} матрицы $\underset{\sim}{\mathbf{c}}$ определяется как

$$c_{ij} = \sum_k a_{ik} b_{kj}$$

Таким образом, чтобы получить элементы c_{ij}, нужно элементы i-й *строки* матрицы $\underset{\sim}{\mathbf{a}}$ попарно перемножить с элементами j-го *столбца* матрицы $\underset{\sim}{\mathbf{b}}$. Например, если

$$\underset{\sim}{\mathbf{a}} = (a_{11}, a_{12})$$

и

$$\underset{\sim}{\mathbf{b}} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

то

$$\underset{\sim}{\mathbf{a}}\,\underset{\sim}{\mathbf{b}} = (a_{11}, a_{12}) \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = (a_{11}b_{11} + a_{12}b_{21}, a_{11}b_{12} + a_{12}b_{22})$$

Таким образом,

$$\underset{\sim}{\mathbf{c}} = (c_{11}, c_{12})$$

где

$$c_{11} = a_{11}b_{11} + a_{12}b_{21}$$

$$c_{12} = a_{11}b_{12} + a_{12}b_{22}$$

Ясно, что перемножать матрицы можно только в том случае, когда число столбцов матрицы $\underset{\sim}{\mathbf{a}}$ равно числу строк матрицы $\underset{\sim}{\mathbf{b}}$. В общем случае матрицу $n \times m$ можно умножить только на матрицу $m \times p$, где n и p произвольны; получающаяся матрица имеет размер $n \times p$. В только что рассмотренном примере для матрицы $\underset{\sim}{\mathbf{a}}$ $n = 1, m = 2$, а для матрицы $\underset{\sim}{\mathbf{b}}$ $m = 2, p = 2$; умножение $\underset{\sim}{\mathbf{a}}$ на $\underset{\sim}{\mathbf{b}}$ дает матрицу $\underset{\sim}{\mathbf{c}}$ размером $n \times p$, или 1×2 (матрица-строка).

Рассмотрим еще один пример. Пусть

$$\underset{\sim}{\mathbf{a}} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

$$\underset{\sim}{\mathbf{b}} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

$$\underset{\sim}{\mathbf{c}} = \underset{\sim}{\mathbf{a}}\,\underset{\sim}{\mathbf{b}} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}$$

Таким образом, если матрицу 2×2 умножить на матрицу 2×2, произведение также будет матрицей 2×2.

Если матрицу 1×2 умножить на матрицу 2×1, получится матрица 1×1, которая является просто числом (скаляром) и не рассматривается более как матрица. Например,

$$\underset{\sim}{\mathbf{a}} = (a_{11}, a_{12})$$

$$\underset{\sim}{\mathbf{b}} = \begin{pmatrix} b_{11} \\ b_{21} \end{pmatrix}$$

$$\underset{\sim}{\mathbf{a}}\,\underset{\sim}{\mathbf{b}} = a_{11}b_{11} + a_{12}b_{21}$$

Если матрица является квадратной (число строк равно числу столбцов), она может быть возведена в любую степень. Например, для квадратной матрицы $\underset{\sim}{\mathbf{M}}$

$$\underset{\sim}{\mathbf{M}}^3 = \underset{\sim}{\mathbf{M}}\,\underset{\sim}{\mathbf{M}}\,\underset{\sim}{\mathbf{M}}$$

$$\underset{\sim}{\mathbf{M}}^N = \prod_{i=1}^{N} \underset{\sim}{\mathbf{M}}$$

(последняя запись означает, что $\underset{\sim}{\mathbf{M}}$ умножается на себя N раз). Эта операция возможна только с квадратной матрицей, потому что только в этом случае выполняется условие, необходимое для перемножения матриц. Отметим также, что если $\underset{\sim}{\mathbf{M}}$ имеет размер $n \times n$, то и $\underset{\sim}{\mathbf{M}}^N$ должна иметь размер $n \times n$.

Нахождение матрицы, обратной данной

Для квадратной матрицы $\underset{\sim}{\mathbf{a}}$ можно найти обратную матрицу $\underset{\sim}{\mathbf{a}}^{-1}$, такую, что

$$\underset{\sim}{\mathbf{a}}^{-1}\underset{\sim}{\mathbf{a}} = \underset{\sim}{\mathbf{I}}$$

где $I_{ij} = 0$ для $i \neq j$; $I_{ij} = 1$ для $i = j$. Матрица $\underset{\sim}{\mathbf{I}}$ называется *единичной*. Например, если $\underset{\sim}{\mathbf{I}}$ — матрица 2×2, то

$$\underset{\sim}{\mathbf{I}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Отметим, что $\underset{\sim}{\mathbf{I}}$ обладает тем свойством, что любая матрица, умноженная на нее, остается неизменной. Таким образом, $\underset{\sim}{\mathbf{I}}$ аналогична числу 1 в скалярной алгебре, так что $\underset{\sim}{\mathbf{a}}\,\underset{\sim}{\mathbf{I}} = \underset{\sim}{\mathbf{a}}$. Найти $\underset{\sim}{\mathbf{a}}^{-1}$ сравнительно просто. Пусть, например,

$$\underset{\sim}{\mathbf{a}} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

Тогда $\underset{\sim}{\mathbf{a}}^{-1}$ находится следующим образом:

$$\underset{\sim}{\mathbf{a}}^{-1} = \begin{pmatrix} a_{22}/\alpha & -a_{12}/\alpha \\ -a_{21}/\alpha & a_{11}/\alpha \end{pmatrix}$$

где

$$\alpha = a_{11}a_{22} - a_{21}a_{12}$$

т.е. α является детерминантом матрицы $\underset{\sim}{\mathbf{a}}$. (Ясно, что для существования $\underset{\sim}{\mathbf{a}}^{-1}$ детерминант не должен быть равен нулю.)

Приведение матрицы к диагональному виду

Для матрицы 2×2 существует матрица $\underset{\sim}{\mathbf{T}}$, такая, что

$$\underset{\sim}{\mathbf{T}}^{-1}\underset{\sim}{\mathbf{a}}\,\underset{\sim}{\mathbf{T}} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

где λ_1 и λ_2 — собственные числа матрицы $\underset{\sim}{\mathbf{a}}$, а $\underset{\sim}{\mathbf{T}}$ называется преобразующей матрицей. Вычислить собственные числа не составляет труда. Пусть

$$\underset{\sim}{\mathbf{a}} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

Чтобы найти λ_1 и λ_2, нужно решить уравнение

$$0 = \begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} = (a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} =$$

$$= \lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - a_{12}a_{21}$$

Используя формулу для нахождения корней квадратного уравнения, получаем

$$\lambda = \{(a_{11} + a_{22}) \pm [(a_{11} + a_{22})^2 - 4(a_{11}a_{22} - a_{12}a_{21})]^{1/2}\}/2$$

Пусть λ_1 соответствует знаку «+», а λ_2 знаку «−», т.е. $\lambda_1 > \lambda_2$. Найдем матрицу $\underset{\sim}{\mathbf{T}}$. Если

$$\underset{\sim}{\mathbf{a}} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

то

$$\underset{\sim}{\mathbf{T}} = \begin{pmatrix} (\lambda_1 - a_{22})/a_{21} & (\lambda_2 - a_{22})/a_{21} \\ 1 & 1 \end{pmatrix}$$

$$\underset{\sim}{\mathbf{T}}^{-1} = \begin{pmatrix} a_{21}/(\lambda_1 - \lambda_2) & (a_{22} - \lambda_2)/(\lambda_1 - \lambda_2) \\ -a_{21}/(\lambda_1 - \lambda_2) & (\lambda_1 - a_{22})/(\lambda_1 - \lambda_2) \end{pmatrix}$$

где λ_1 и λ_2 — собственные числа матрицы $\underset{\sim}{\mathbf{a}}$.

Укажем два других полезных свойства матрицы $\underset{\sim}{\mathbf{a}}$ размером 2×2.

$$\lambda_1 + \lambda_2 = a_{11} + a_{22}$$

$$\lambda_1\lambda_2 = a_{11}a_{22} - a_{12}a_{21}$$

Этих соотношений вместе с только что приведенными выражениями для $\underset{\sim}{\mathbf{T}}$ и $\underset{\sim}{\mathbf{T}}^{-1}$ достаточно, чтобы получить выражения для преобразующей матрицы в уравнении (20.53) и обратной ей в уравнении (20.54). Нетрудно также обобщить эти результаты на случай матриц $n \times n$, где $n > 2$.

Полезное введение в матричную алгебру можно найти в гл. 1 книги «Методы прикладной математики» Ф. Хильдебранда [F. B. Hildebrand, Methods of Applied Mathematics (Englewood Cliffs, N. J.: Prentice-Hall, 1965)].

Приложение Б

Решение задач

ГЛАВА 1

1.1. Диаметр соответствующего цилиндра равеи $2\pi r$. После проецирования один виток спирали будет соответствовать гипотенузе прямоугольного треугольника, катеты которого равны $2\pi r$ и P.

1.2. Регуляторный механизм поддерживает коицентрацию белка на достаточно высоком уровне, чтобы мог образовываться стабильный октамер, не допуская перепроизводства.

ГЛАВА 2

2.1. Наличие пяти полос, наблюдаемых для смеси B + C, означает, что белки являются тетрамерами. Полосы соответствуют B_4, B_3C, B_2C_2, BC_3 и C_4. Для объяснения наличия только трех полос в смесях A + C или A + B предположим, что димер A_2 очень стабилен по сравнению с гетеродимерами AC и AB. Изолированный A будет тетрамером $(A_2)_2$. Три полосы смеси A + B соответствуют $(A_2)_2$, A_2B_2 и B_4. Аналогичиые полосы можно увидеть для смесей A + C.

2.2. При добавлении мономера к димеру формируются два контакта белок — белок, тогда как при образовании димера из двух мономеров или тетрамера из тримера и мономера — только один. Сходным образом при добавлении последнего мономера, завершающего процесс образования диска, формируются четыре коитакта белок — белок, а при добавлении мономера к менее завершениому диску или к полностью завершенному число контактов будет меньше. Если предположить, что при формировании любых коитактов выигрыш в энергии одинаков, то оказывается, что тримеры (пентамеры, гептамеры и т.д.) и диски будут предпочтительными по сравнению с другими олигомерами с бóльшим или меньшим числом мономеров.

2.3. Первичная структура, вероятно, представлена двумя почти одинаковыми участками последовательности (возможно, она подобна структуре ферредоксина). N- и C-концы, по-видимому, расположены в третичной структуре очень близко друг от друга.

2.4. α-Спираль, состоящая из 85 остатков, имела бы длину 128 Å. Таким образом, рассматриваемый белок не может находиться в строго α-спиральной конформации. Цепь из 85 остатков в конформации β-слоя была бы растянута до 295 Å. Если она складывается в двухцепочечный антипараллельный β-слой, результирующая структура будет иметь длину около 150 Å. Можно также сформировать структуры такой длины из различных комбинаций спиралей и слоев. Однако в любой из этих структур большая часть боковых групп аминокислот (или все они) была бы экспонирована. Если белок содержит обычный набор гидрофобиых остатков, то весьма маловероятно, что какая-нибудь из этих структур будет стабильна в водном растворе. В рибосоме такие структуры могут стабилизироваться вследствие образования множества контактов боковых цепей с другими белками (или с рибосомальной РНК). Если поместить такой белок в раствор, он почти наверное примет более компактную конформацию за счет разрушения некоторых слоев или спиралей.

2.5. Обратите внимание на то, как чередуются полярные и неполярные остатки в последовательности. Если такой белок сложен в двухцепочечный антипараллельный β-слой, то одна его поверхность будет полностью полярной, а другая—неполярной. Такой мономер будет легко прилипать к мембране своей неполярной поверхностью. Он также будет легко димеризоваться в результате гидрофобных взаимодействий, образуя структуру типа сэндвича с неполярной внутренней областью.

ГЛАВА 3

3.1. Вероятность обнаружения заданной последовательности из следующих друг за другом оснований в данном месте молекулы равна $1/4^n$. В ДНК из 10^4 оснований эта последовательность встретится в среднем $10^4/4^n$ раз. Следует выбрать n таким образом, чтобы эта величина была меньше или равна 0,01. Это дает $0{,}01 \approx 10^4/4^n$, откуда $n \approx 10$.

3.2. В В-форме ДНК длина двойной спирали составит $13{,}1 \cdot 10^{-3}$ см, что соответствует $3{,}79 \times \times 10^4$ виткам. В А-форме ДНК длина составит $9{,}70 \cdot 10^{-3}$ см, а число витков — $3{,}44 \cdot 10^4$.

3.3. Эта проблема обсуждается в работе U. H. Settler et al., J. Mol. Biol., **131**, 21 (1979)[1].

3.4. Одна половина молекулы образует одноцепочечное кольцо, а вторая закручена вокруг первой, образуя двойную спираль. Авторам неизвестно, была ли такая структура обнаружена экспериментально.

3.5. В ДНК тоже будут образовываться спиральные шпильки, однако их структура будет, вероятно, такой же, как в В-форме ДНК, а не как в РНК-11. В то же время часть третичных взаимодействий сохранится, а некоторые из них станут невозможными. К последним относятся, например, водородные связи, в образовании которых принимает участие 2′-ОН-группа рибозы. Взаимодействия с участием модифицированных нуклеотидов также окажутся невозможными для ДНК.

ГЛАВА 4

4.1. См. рис. 9 в работе: C. Sardet, A. Tardieu, V. Luzzati, J. Mol. Biol., **105**, 383 (1976).

4.2. Это нелегкая задача. Можно попытаться предотвратить случайные столкновения с помощью образования поперечных сшивок в мембране при температурах ниже фазового перехода (переходов) в липидах. Если есть возможность проверить активность белков, можно попытаться восстановить очищенные сшитые белковые пары и измерить активность. Наиболее эффективная мыслимая стратегия — путем разбавления уменьшить концентрацию белков и посмотреть, образуются ли поперечные связи и в этом случае, но сделать это весьма непросто. Дальнейшее обсуждение этого вопроса можно найти в статье: K. Peters, F. M. Richards, Ann, Rev. Biochem., **46**, 523 (1977).

4.3. Возможно хорошее соответствие C_2-осей симметрии белка и нуклеиновой кислоты.

4.4. Нарушение периодичности дестабилизирует высокоупорядоченные спиральные и олигомерные структуры, и тогда единственное, что может образоваться, — это менее упорядоченная структура геля.

4.5. Соли должны дестабилизировать комплексы между нуклеиновыми кислотами и основными (положительно заряженными) белками. Эти комплексы связывают меньше противоионов, чем отдельные компоненты. Напротив, комплексы между нуклеиновыми кислотами и кислотными (отрицательно заряженными) белками должны стабилизироваться солями. Эти комплексы будут связывать больше катионов, чем отдельные компоненты. Комплексы такого типа обнаружены в нуклеопротеидах, выделяемых из галофильных организмов. Вероятно, они являются результатом адаптации этих организмов к высокой внутриклеточной концентрации солей.

ГЛАВА 5

5.1. Напомним, что все остатки, предшествующие пролину, имеют измененную конформационную энергетическую карту, за исключением глицина, на который следующий за ним пролин влияет не очень сильно. (Для случая остатка аланина, предшествующего пролину, см. рис. 5.11.) В последовательности Ala-Pro область 1 стерически запрещена для остатка аланина, а именно она соответствует менее растянутой конформации, которая может быть принята при образовании изгиба. Следовательно, для образования изгиба, что необходимо при формировании циклического пептида, лучше, чтобы перед пролином находился не аланин, а глицин. Из-за отсутствия боковой цепи остаток гли-

[1] Рассмотрение этого вопроса невозможно без учета образования левоспиральной Z-формы [см. F. M. Pohl, R. Thomae, E. DiCapua, Nature, **300**, 545 (1982)]. — *Прим. ред.*

цина может принять все доступные ему менее растянутые конформации даже тогда, когда за ним следует пролин. Поэтому предпочтительной последовательностью является следующая: X = Ala, Y = Gly, Z = Pro.

5.2. Для решения этой задачи следует использовать табл. 5.8 и 5.9. Назовем конформацию, в которой все три остатка расположены во внутренней области белка, конформацией A, а конформацию, в которой они экспонированы, — конформацией B. Из табл. 5.8 следует, что $\Delta G_{\text{пер}} = 0{,}73$, 2,42 и 2,65 ккал/моль для Ala, Leu и Phe соответственно. Сумма этих свободных энергий $\Delta G_{AB} = 5{,}80$ ккал/моль и в первом приближении равна свободной энергии переноса из A в B в воде. Следовательно, $K = (A)/(B) = \exp(-\Delta G_{AB}/RT) = 1{,}8 \cdot 10^4$. Согласно табл. 5.9 свободная энергия переноса Ala, Leu и Phe в водный раствор мочевины уменьшается соответственно до 0,07, 0,38 и 0,70 ккал/моль. Таким образом, ΔG_{AB} уменьшается до 1,15 ккал/моль и, следовательно, для конформационного изменения в мочевине $K = 2{,}6 \cdot 10^3$.

5.3. Применим подход, использованный в разд. 5.8. Пусть ΔG_{AB} — свободная энергия перехода из конформации A в конформацию B. Следовательно,

$$\Delta G = -\int_{r_2}^{r_1} (F)dr = \gamma(r_2^{-1} - r_1^{-1}), \text{ где } \gamma = aT$$

$$K = \exp(-\Delta G/RT) = \exp[-\gamma(r_2^{-1} - r_1^{-1})/RT]$$

$$\Delta H = -T^2 d(\Delta G/T)/dT = -T^2 d[a(r_2^{-1} - r_1^{-1})]/dT = 0$$

$$\Delta S = (\Delta H - \Delta G)/T = \gamma(r_1^{-1} - r_2^{-1})/T$$

5.4. Обратимся к уравнениям (5.13) и (5.14). Так как $\chi = 0{,}1$ и в этаноле, и в воде, вклад энтропии смешения в свободную энергию переноса $\Delta G_{\text{сумм}}$ равен нулю. Однако вклад энтропии смешения в $\Delta G^*_{\text{сумм}}$ отличен от нуля, так как $\chi = 1$ в этаноле и $\chi = 0{,}1$ в воде. Поэтому $\Delta G^{0*}_{\text{сумм}} = \Delta G^0_{\text{сумм}} + RT\ln 0{,}1$ и $\Delta G^{0*}_{\text{сумм}} - \Delta G^0_{\text{сумм}} = RT\ln 0{,}1$. Следовательно, экспериментатор неправ.

5.5. Из разд. 5.4, а также и из рис. 5.7 и 5.8 мы знаем, что наличие остатков глицина в цепи способствует принятию более компактных, менее растянутых конформаций. Действительно, для одного и того же числа остатков среднеквадратичное расстояние между концами полиглицинового статистического клубка приблизительно вдвое меньше, чем для полиаланинового (см. разд. 5.4, а также гл. 18). Поэтому если нерастянутые волокна имеют одинаковую длину, волокно, содержащее полиглицин, должно содержать вдвое большее число остатков. При растяжении цепь стремится принять растянутую конформацию типа зигзага (*транс*), в которой $\phi = \psi = 180°$. Волокно, содержащее полиглицин, будет растянуто сильнее, так как, во-первых, составляющие его цепи в нерастянутом состоянии более компактны, а во-вторых, в этих цепях находится больше остатков. Следовательно, волокно 1 содержит полиглицин, а волокно 2 — полиаланин.

ГЛАВА 6

6.1. Уравнение (6.6) для димеризации получается непосредственно из определения K_2. Используя обозначения и определения, введенные в разд. 6.7, нужно выразить K_2 через m' и ϕ_p. Поскольку $m_2 = m' - m$, $m_1 = 2m - m'$ и $m = \phi_p m'$, то

$$K_2 = m_2/m_1^2 = (m' - m)/(2m - m')^2 = m'(1 - \phi_p)/m'^2(2\phi_p - 1),$$

или

$$K_2 m' = (1 - \phi_p)/(2\phi_p - 1).$$

Это и есть уравнение (6.6).

Чтобы вывести уравнение (6.7), следует начать с определения ϕ_p, имея в виду, что $K = m_2/m_1^2 = m_3/(m_2 \cdot m_1) = \ldots$.

$$\begin{aligned}\phi_p &= (m_1 + m_2 + \ldots + m_n + \ldots)/(m_1 + 2m_2 + \ldots + nm_n + \ldots) = \\ &= (1 + Km_1 + Km_2 + \ldots)/(1 + 2Km_1 + 3Km_2 + \ldots) = \\ &= (1 + Km)/(1 + Km' + Km) = \\ &= (1 + Km'\phi_p)/(1 + Km' + Km'\phi_p).\end{aligned}$$

Данное соотношение упрощается до

$$Km' = (1 - \phi_p)/\phi_p^2.$$

Это и есть уравнение (6.7).

6.2. Раствор I: взанмодействие между основаниями отсутствует.

Раствор II: происходит димеризация, вероятно, за счет взаимодействия, сходного с тем, которое имеет место при образовании водородных связей; оно препятствует образованию более крупных, чем димеры, комплексов. Если $2A \rightleftarrows A_2$ и $K = (A_2)/(A)^2$, то

$$(A)_0 = (A) + 2(A_2) = (A) + 2K(A)^2 = (A) + C(A)^2,$$

где $C = 2K$.

Раствор III: происходит неограниченная ассоциация мономеров, причем константа ассоциации одинакова для всех этапов. Поэтому если $K = (A_2)/(A)^2 = (A_3)/(A_2)(A) = \ldots$, то

$$(A)_0 = (A) + 2(A_2) + 3(A_3) + \ldots = (A) + 2K(A)^2 + 3K^2(A)^3 + \ldots = C_1(A) + C_2(A)^2 + C_3(A)^3 + \ldots,$$

где $C_1 = 1$, $C_2 = 2K$, $C_3 = 3K^2$, $C_4 = 4K^3$. Таким образом,

$$C_2 = 2K, \qquad C_k = (k/2^{k-1})C_2^{k-1}.$$

В основе неограниченной ассоциации скорее всего лежит стэкинг оснований.

6.3. Различить две рассматриваемые ситуации позволяют уравнения (6.6) и (6.7), так что, вероятно, в ходе эксперимента допущена ошибка. Следует отметить, однако, что при $\phi_p \approx 1$ левые части уравнений (6.6) и (6.7) равны. В этой области величина $1 - \phi_p$ очень мала, и левые части обоих уравнений оказываются близкими к нулю. Это означает, что $m' \ll K_2^{-1}$ [уравнение (6.6)] или $m' \ll K^{-1}$ [уравнение (6.7)], т.е. концентрация мононуклеотидов слишком мала для того, чтобы наблюдать хоть какую-то ассоциацию. Это соображение, однако, должно быть очевидно для экспериментатора; в противном случае оно могло бы объяснить результаты эксперимента.

6.4. Структурные переходы в ДНК более подробно рассмотрены в гл. 22 и 23. Можно ожидать, что параметры кривых плавления будут зависеть от нуклеотидного состава ДНК, так как GC-пары стабильнее AT-пар, а стэкинг-взанмодействие, которое стабилизирует спираль, будет зависеть от нуклеотидного состава и даже от последовательности. Поскольку образование двойной спирали сопровождается объединением одноименно заряженных нитей, на плавление ДНК будет оказывать существенное влияние и концентрация соли в растворе.

Кооперативность будет наблюдаться в том случае, когда переход в новое состояние первой пары оснований осуществляется с бо́льшим трудом, чем переход следующих пар. Помимо того что при образовании первой пары (нуклеации) имеется энтропийный барьер, связанный с необходимостью расположения двух цепей надлежащим для спаривания образом, эта пара лишена возможности участвовать в стэкинг-взанмодействии. Однако образование рядом последующих пар уже сопровождается стэкинг-взанмодействием с ранее образовавшимися парами, так что увеличение длины спиральной области выгоднее, чем нуклеация, в результате чего и возникает кооперативность.

6.5. Из гл. 5 известно, что мочевина увеличивает растворимость в водных растворах аминокислот с неполярными боковыми группами. Отсюда можно заключить, что она будет разрушать гидрофобные взаимодействия. Поэтому в той степени, в какой стэкинг оснований стабилизирован гидрофобными взаимодействиями, мочевина должна дестабилизировать стэкинг в одноцепочечных структурах и в двойных спиралях нуклеиновых кислот, чему имеются экспериментальные подтверждения.

Предметный указатель

Агрегация 83, 85—87, 93
Аденилаткиназа 71—72, 120, 281—283
Аденовирус 207
Аденозинтрифосфатаза $Na^+ — K^+$ 58
Активный центр лизоцима 35
— — сериновых протеаз 84
— — субтилизина 120
— — химотрипсина 30, 120
Актин (актиновые полимеры) 139
Актиномицин 310
Алкогольдегидрогеназа 92, 117, 119
Аллоизолейцин 43
Аллотреонин 43
Амидная группа 239, 245
— связь 239
Аминоацилирование 155
Аминокислотная последовательность 71, 73, 82—83
— — родственных белков 77
Аминокислотный состав 54
Аминокислоты 43—44
— доступность в белках 113
— ионизация боковых групп 50—51
— классификация свойств 53
— конформационные параметры 278—281
— положение боковых групп в белках 100—101, 106—107
— полярность боковых групп 52
— структура 44
— титрование 48—52
— D 43
— pK_a 50—51
Амфифильные молекулы 214, 218—219
Антигены 227—228
Антикодоновая петля тРНК 180—181
Антитела 27, 223, 227—228
Аполипопротеиды 232—233
Апофермент 62
Апоферритин 127
Асимметрия бислоев 219—220
— липидного состава эритроцитов 230
— мембран 228. *См. также* Ферритин
Аспартат-транскарбамоилаза (АТСаза) 55, 133—138
Ассоциация субъединиц 135, 137—139, 141
N-ацетилмурамовая кислота 201
N-ацетилглюкозамин 201
Ацетилхолина рецептор 58
Ацетилхолинэстераза 58

Бактериофаги 23, 139, 143, 165, 210
Белки, аминокислотная последовательность 71, 73, 75, 82—83
— агрегаты типа «завиток» 144
— — трубчатые 83
— взаимодействие с липидами 230
— — — нуклеиновыми кислотами 207—209, 211—212
— внутренние молекулы воды 109
— вспомогательные (при сборке вирусов) 31, 210
— гидрофобность, гидрофобные остатки 56, 100, 107, 223
— динамика структуры 114

Белки, длина цепи 65
— домены 67, 102, 106, 118—119
— заряд 28, 107, 135. *См. также* Дискриминантная функция *Z*
— ионизационное равновесие 48
— конформационный анализ 237
— кооперативная полимеризация 141
— мембранные 57—58, 73, 107, 221, 223—224, 229—231
— многоцепочечные 60
— объем молекулы 111
— плотность 111
— проницаемость для растворителя 113
— распределение в мембранах 226
— рибосомные 211—212
— спиральные четвертичные структуры 86, 139—140
— структура 261, 267
— титрование 48—51
— трансмембранные 223, 230
— третичная структура, методы сравнения 114—115
— укладка молекулы 71—73
— эволюция 28—29, 78, 120—121
Белково-нуклеиновая организация 211—212. *См. также* Белки, взаимодействие с нуклеиновыми кислотами
Белок пурпурной мембраны 55, 58
Биотин 64
Бислои 202, 218—221, 230
Бислой плоский 218—219
— сферический 219
β -Бочонок 121
Бромелаин 206

Вандерваальсовы контакты 129, 138
Везикулы *см.* Пузырьки
Винтовая ось симметрии 14, 125, 130
Вирус табачной мозаики 20, 144—145, 207, 209—210
Вирусы 20, 143—145, 209
— сборка 144—145

Влажность, влияние на структуру ДНК 173
Вода 109—111, 261—263, 265—266
Водородные связи 35, 89, 91, 129, 179—183, 198, 208—257
Вращение белков в мембранах 228—229
Вставки в нуклеиновых кислотах 157, 165
Вторичная структура 15, 87—96, 160, 167, 232

β -Галактозидазы ген 159
Гексокиназа 98, 130
Гели 197, 198, 200
Гемоглобин 18, 20, 24, 34, 36—37, 49, 62, 64, 77—87, 108, 127, 129, 132—133
— агрегация 83—87
— инвариантные остатки 80
— конформационные изменения 18—19
— линейный дихроизм 83
— последовательность 80—81
— предсказание структуры 282
— псевдосимметрия 20
— связывание дифосфоглицерата 24 36—37
— — кислорода 36
— титрование 49
— четвертичная структура 18—19
— эволюция 77—78
— Сидней 39
— S (дезоксигемоглобин S) 83—87
Гемоглобины мутантные 39—40, 80—81
Генетическая карта 167
Генетический код 158
Гетеродуплексное картирование 167
Гетеродуплексы 164—165
Гиалуроновая кислота 198—199
Гибкость структуры 27, 109, 113—114, 190

Гибридизация *см.* Гибриды
Гибридные структуры 131—132
Гибриды (гибридные ДНК, гибриды ДНК—РНК) 156, 164—167, 170, 175
п-Гидроксимеркурибензоат 134
Гидроксипролин 95
Гидролиз РНК щелочной 156
Гидрофобное связывание 267
Гидрофобность 56—57
Гидрофобные взаимодействия 139, 267, 270, 309—310
— силы 54
Гистерезис при титровании 49
Гистонов гены 166
Гистоны 55, 73, 75, 212
Гликоген 197
Гликозидная связь 289
Гликолипиды 216, 221
Гликопротеиды 73, 204, 207, 221, 230. *См. также* Гликофорин
Гликофорин 73, 76, 204—205, 230
Глицеральдегидфосфатдегидрогеназа 117, 119, 139, 230
Глицин 93
Глобины 77—78
Глутаминовая тРНК (тРНК^{Glu}) 161
Глюкоза 98
Глюкуроновая кислота 197
Гомология 164
— первичных структур белков 119. *См. также* Аминокислотная последовательность, сходство различных белков
Гормоны 37
Грама реакция 201
Гуанидингидрохлорид 135

Двойная спираль 168—169, 174
Двухвалентные катионы 33. *См. также* по названиям
Дезоксигемоглобин S 83—87
Делеции в ДНК 165
Денатурация 30, 164
Десмозин 68
Дигидроуридиловая петля в тРНК 179
Дигидроуридин 186
Диметилсуберимидат 131
Динамика структуры белков 113
Дипольные взаимодействия 245—247
Дискриминантная функция Z 57—58
Дисперсионная энергия *Лондона* 244
Диссоциация субъединиц 131—132, 134—138
Дисульфидные связи 66, 79, 82, 93, 270
Дисульфидный обмен 66
Дифосфоглицерат 24, 36—37
1,5-Дифтор-2, 4-динитробензол 70, 233
Диэдрическая симметрия 126
ДНК 171
— антипараллельность цепей 172—173
— взаимодействие с гистонами 75
— вставки 165
— гибкость 190
— двойная спираль 171, 174
— делеции 165
— бактериофага T2 22—23
— — T5 156
— — ϕ X174 154
— желобки 173—174, 208—209
— изломы 188, 190—192
— инверсии 165
— конкатемерная 210
— контурная длина 42, 167
— конформации (A- и B-формы) 176—177
— — пиранозного кольца 176
— направление цепей 170. *См. также* ДНК, антипараллельность цепей
— нуклеотидный состав 153
— последовательность *см.* Нуклеотидные последовательности
— — *lac*-оперона 159
— сдвиг равновесной конформации двойной спирали 176
— структура двойной спирали 170—172

ДНК, стэкинг оснований 173
— третичная структура 190
— тройная спираль 177—178
— SV40 193
ДНК-полимераза 172
Домены в структуре белков 67, 102, 106, 118

Escherichia coli 21, 122, 160—161, 187, 201, 207, 218

Желобки в молекуле ДНК 173—174, 208—209
Жидкостно-мозаичная модель мембраны 228
Жирные кислоты 214—215

Замораживание—скалывание 224—225
— травление 224—226
Заряд белков 28, 107. *См. также* Дискриминантная функция Z
Заряда перенос *см.* Системы с переносом заряда
Заряженные остатки (группы) 75, 100, 106
Зеркальное отражение 124—125

β-Изгибы 71—72, 91, 95, 281
Изломы 188, 190—192
С-55-Изопреноид-алкоголь — фосфокиназа 231
Икосаэдрическая симметрия 123, 126
Иммуноглобулин G (IgG) 27, 66—67, 121
Инвариантные остатки 80
Инверсии (в ДНК) 165
Инверсионная ось симметрии 125
Ингибирование по механизму обратной связи 134
Инициации участки 163—164
Инициирующие тРНК 160, 162
Инозин, образование нуклеотидных пар 163
Инсулин 31—32, 129
Интегральные мембранные белки 222, 230
Интеркаляция оснований в тРНК 186
Ионизации свободная энергия 46—47
Ионизация боковых групп аминокислот 49—51
Ионная атмосфера 48
Ионизационное равновесие 46—51, 149
Ионные взаимодействия 268. *См. также* Солевые мостики

Кальций (Ca^{2+}) 102
Каналы в мембранах 224
Карбоксилат-аниона комплекс с гуанином 209
Карбоксипептидаза 62, 112
Кардиолипин 217
Каррагенаны 198
Картирование генов 165—167
Катенан 157
Кислорода связывание гемоглобином 36
Кластеры антигенов на поверхности мембраны 227—228
«Клеверный лист» 160—161, 181
Клеточная стенка 201
— — биосинтез 162
Колицин 212
Коллаген 16, 18, 55, 68, 70, 74, 92—94, 241
Комплементарность 192
Конканавалин А 129, 282
Контурная длина ДНК 42, 167
Конфигурация, определение 12
Конформационная энергия, минимизация 73
Конформационные изменения 18—19, 98
— карты 249—254, 291—293

Конформационные параметры аминокислот 278—281
Конформационный анализ 237
Конформация, определение 11—12, 18
— сахара в РНК 175
Кооперативная полимеризация белков 141
Кофакторы 65. *См. также* Металлы, Простетические группы
Критическая концентрация 143
Ксилан 198
Кубическая симметрия 123, 126—127

Лактатдегидрогеназа 116—129, 126, 131
β -Лактоглобулин 51
Латеральное перемещение белков в мембране 228—229
— фазовое разделение белков в мембране 220, 221
Лецитины 214
Лиганды, связывание 33—34
Лизилоксидаза 70
Лизин 47, 51, 70
Лизоцим 34—35, 51, 55, 96, 100—101, 109, 112, 275, 277
Линейный дихроизм 83
Липидные пузырьки 218. *См. также* Бислой сферический, Липосомы
Липидный бислой 202, 218—221
— — пограничные молекулы 230
— состав мембран 218
Липидов связывание 233
Липиды, взаимодействие с белками 230
— подвижность в мембранах 230
— свойства 214
— фазовые переходы 221
Липоевая кислота 64
Липополисахарид клеточной стенки бактерий 201
Липопротеиды 201—202, 231—232
Липосомы *см.* Бислой сферический, Липидные пузырьки
Липотейхоевая кислота 201, 204

Малатдегидрогеназа 117, 119, 126
Манноза 204—205
Матричная алгебра 315
Матричные методы многокомпонентного анализа 41
Мембрана митохондриальная 218, 222
— плазматическая 213, 222, 229
— эритроцита 218, 229
Мембранные белки 57—58, 73, 107, 221, 223—224, 229—231
— — вращение 228—229
— — интегральные 222, 230
— — латеральное перемещение 228—229
— — олигомеры 224
— — периферические 221
Мембраны 213, 218, 221—222, 228
— белковые компоненты 221
— жидкостно-мозаичная модель 228
— липидный состав 218
— метод поверхностной метки 224
— температура фазового перехода 226, 231
— трансмембранные белки 223, 230
— электронная микроскопия 224
Металлопротеиды 65
Металлы 60—62
Миелин 218
Микротрубочки 139
Миоглобин 16, 18, 51, 62, 77, 88, 109, 275—276, 284
Митохондриальная мембрана 218, 222
Мини-петля в тРНК 179
Многокомпонентный анализ 41
Модели двойных спиралей 174
— мембраны 228
— молекул 283
— третичной структуры 96—103
Модель агрегации субъединиц 145
Модельные системы 38
Модификация химическая 40, 113, 160, 162, 186—188

Молекулярная динамика 114
— масса белков 126—127, 130
— — липопротеидов 232
— — нуклеиновых кислот *см.* Нуклеиновые кислоты, длина молекулы
— — различных биополимеров 10, 12
Молекулярные модели, построенные с помощью ЭВМ 283
Молекулярный объем белков 111
Мононуклеозиды 306
мРНК 163—164
Мутантные формы 24, 31, 33, 39—40, 80—81
Муцины 206

Натрия ионы (Na^+) 174
Невалентные взаимодействия 244
Необычные основания 149, 153, 179
Неспаренные основания 186—187
Нуклеазы рестрикции 158
Нуклеация 86, 143
Нуклеиновые кислоты 287
— — взаимодействие с белками 207—209
— — геометрия 289
— — длина молекулы 156
— — и полиамины 155
— — конформационный анализ 287—293
— — нуклеотидный состав 153—154
— — одноцепочечные 186—187
— — последовательность оснований 160—161, 164
— — — палиндромов 159
— — силы, стабилизирующие структуру 204—311
— — спаривание оснований 295
— — третичная структура 311
— — углы вращения остова, стерические ограничения 289—290
Нуклеозидов таутомерные превращения 150, 152
Нуклеопротеидные комплексы 207—209
Нуклеопротеиды 207, 210—211
Нуклеосомы 207, 212—213
Нуклеотидные последовательности, анализ частоты встречаемости различных комбинаций звеньев 172
Нуклеотидный состав ДНК 153
— — РНК 154
Нуклеотидсвязывающие домены в белках 116—117
Нуклеотиды 148
NAD 116—119

Обмен дисульфидный 66
Объем молекулы белка 110
Овальбумин 206
«Окрашивание» антител 226—228
Октаэдрическая симметрия 123, 126
Олеиновая кислота 215
Олигомеры мембранных белков 224
Олигопептиды 44, 46
Олигосахариды 60
Опухолеродные вирусы 193
Осмотический коэффициент эффективный 307
Основания (нуклеозиды) необычные 149, 153, 179
— неспаренные 186—187
Осно́вные белки 207

Палиндромные последовательности 159
Пальмитиновая кислота 215
Пары оснований в кристаллах динуклеотидов 297—298
— — — растворе 298—301
— — термодинамические параметры 299—302
— — угол наклона к оси спирали 168, 171, 176
— — уотсон-криковские 169
— — хугстеновские 295—296
— — — обратные 296
Параллельный перенос 125
Пектины 197—198

Пептидная связь 89, 90
Пептидных связей разрыв 59, 99
Пептидогликаны 162, 201
Первичная структура, определение 13
— — различных белков 73—86
Перенос заряда 29
Периплазматическое пространство 201
Периферические белки мембран 221
D-Петля 157
Пиранозное кольцо 175
Пиридоксальфосфат 64
Пируватдегидрогеназный комплекс 20-21, 127
Пируватоксидаза 231
Плазмалоген 216
Плазматическая мембрана 213, 222, 229
Плотность белков 110
— липопротеидов 232
— упаковки аминокислотных остатков 109—110, 138
Поверхностной метки метод 224
Повторы 72
Полиаденциловая кислота [poly (A)]
— — гипохромизм 304—305
— — спектр поглощения 304
— — стэкинг оснований 304
Полиамины 155
Полиглицин 241—242
Полимеризация 140—143
— кооперативная 141
— матричный механизм 144—145
Полинуклеотидные цепи 237
Полинуклеотиды 170, 177, 288—289
Полипептидные цепи, конформация, геометрия 237—243
— — плоский зигзаг 237
— — свертывание 121
Поли (L-пролин) 241
Полипролиновые спирали 92
Полирибоаденциловая кислота 290
Полисахариды 195, 200
— образование гелей 198
Полярность боковых групп аминокислот 52
— цепей ДНК *см.* ДНК, направление цепей
Поперечные связи 66. *См. также* Дисульфидные связи, Поперечные сшивки
Порфириновое кольцо 62
Последовательность (и) 13, 162, 177
— аминокислотная 71, 73—77, 83
— нуклеотидная 172. *См. также* Нуклеиновые кислоты, последовательность оснований
— сравнение 77, 164
Посттранскрипционная модификация РНК 154
Посттрансляционные модификации 59—60
Потенциальная энергия, ее оценка 243—253
Потенциометрическое титрование 43—45, 49—51
Предсказание структуры 25—26, 70—72, 160, 190, 277
Проинсулин 31
Пролин 52, 92, 109, 251—253
Проницаемость белков для растворителя 113
Простетические группы 62, 64
Протеазы 84. *См. также* Эстеразы
— сериновые 121
Протеолиз *см.* Пептидных связей разрыв
Протонированные полинуклеотидные спирали 170, 177
Псевдосимметрия 20, 158, 168—170, 172
Псевдоуридиловая петля в тРНК 179
Пузырьки 218, 223
Пурпурная мембрана 222. *См. также* Белок пурпурной мембраны
pK_a 46—51
poly (I) 177

Разветвленные цепи 157. *См. также* Гликоген

Рамачандрана карты 241
Рауля закон 308
Регуляция ферментативной активности 36—37
Ренатурация 31, 164—165
— гетеродуплексов 164
Рентгеноструктурный анализ 87—88, 175
lac-Репрессор 127, 159
Рестрикции нуклеазы 158
Рибоза 175—176, 179
Рибонуклеаза 51, 71, 97, 109—111, 204, 271
Рибосомные белки 211—212
Рибосомы 20—22, 160, 207, 211—212
Роданаза 117, 120
Родопсин 58
РНК, взаимодействие с белками 211—212
— вируса табачной мозаики 209—210
— информационная *см.* мРНК
— матричная *см.* мРНК
— нуклеотидный состав 154
— опухолеродных вирусов 193
— посттранскрипционная модификация 154
— рибосомиая *см.* рРНК
— транспортная *см.* тРНК
— щелочной гидролиз 156
РНК-полимераза 160
рРНК 21, 163, 190, 211—212

Самосборка 144, 145 *См. также* Сборка вирусов
Саркоплазматический ретикулум 222
Сателлитная ДНК 153
Сахара в белках 60, 204, 206—207. *См. также* Углеводы
— — гликопротеидах 75, 205, 207
— — муцинах 206
— — нуклеиновых кислотах 175
— — полисахаридах 195, 197, 206
— конформация 170, 175
Сборка 31, 143—145. *См. также* Агрегация, Полимеризация
— вирусов 31, 210. *См. также* Самосборка
Сведберга единицы 20
Свертывание полипептидной цепи 121
Сверхспирали 16, 190—192
Свободная энергия ассоциации 141
— — ионизации 46—47
— — переноса ($\Delta G_{\text{пер}}$) 53
— — при сопряжении реакций 46—47
— — стэкинга 309
Серин, фосфорилирование 59
Сериновые эстеразы 28, 79, 82—83
Серповидноклеточная анемия 86—87
Серповидные клетки 87
Сиаловая кислота 205
Симметрии винтовая ось 14, 125, 130
— инверсионная ось 125
Симметрия 20, 122—124, 137—138, 172, 177
— аксиальная 125
— вращения 122—124
— диэдрическая 126
— кубическая 123, 126—128
— икосаэдрическая 123, 126
— октаэдрическая 123, 126
— тетраэдрическая 126
— пространственная 125
— точечная 122—125
— C_2 (ось симметрии второго порядка) 20, 91, 124—125, 133, 157, 208, 213
Системы с переносом заряда 29
Складчатый слой 240. *См. также* β-Слой
β-Слой (и) 71—73, 90—92, 95, 103—105, 116, 139, 209, 281
Сопряжение 44, 49—51
— реакций 46—47
Сопряжения энергия 46—47
Солевые мостики 59
Спектрин 221, 229
Спираль (и) 14, 170
— двойные 170
— — параметры 172
— геометрия 14
— математическое описание 14

Спирали полинуклеотидные 170, 177
— полисахаридов 196
— радиус 14
— шаг 14
— *n*-го порядка 14
α -Спираль 11, 72—73, 88—91, 116, 209, 231—233, 241—242, 281
D-спираль (D-петля) в тРНК 179—181
Спиральные четвертичные структуры 139
— — — определение длины 143
Стеариновая кислота 215
Структура белков, сходство 116—122
— получение данных 274
Структурные карты 116—118. *См. также* Третичная структура
Структуры предсказание 25—26, 70—72, 160, 190, 277
— семейства 28—29
— уровни 12, 41
Сшивание рибосомных белков 212
— субъединиц 131
Сшивающие агенты 132—133
Сшивки 68, 70, 131—133, 201, 212, 229
Субтилизин 117, 120—121
Субъединичная структура (организация) 122, 126—128, 133
Субъединицы 20, 32—33, 80, 91, 122, 126—127, 129—135, 138—139, 145
Супероксиддисмутаза 121, 126
Стэкинг оснований в двойных спиралях 173
— — — одноцепочечной poly (A) в растворе 304
— — — тРНК 186
— — константа ассоциации 308
— — пуриновый и пиримидиновый 308
— — роль растворителя 309
— — свободная энергия 308
— — температурная зависимость 304—306
— — термодинамика 306—311
Сферический бислой 219
Сфингозин 216
Сывороточный альбумин 51, 66
Staphylococcus aureus 201

Таутомерные формы 150
Текучесть бислоев *см.* Фазовый переход в липидах
Температура, ее влияние на свойства липидов 221
— перехода 221, 226, 231
Термодинамика процессов растворения 52—53
— стэкинга 306—311
Термолизин 102, 282
Тетраэдрическая симметрия 123, 126
Тимина димеры 309—310
4-Тиоуридин (тио-U) 187—188
Тирозиновый супрессор тРНК 161
Титрование гемоглобина 49
— нуклеозидов и нуклеотидов 149—150
— потенциометрическое 43—44, 49—51
Торсионные углы *см.* Углы внутреннего вращения
Торсионный потенциал 247
Точечная симметрия 122—129
Транскрипции регуляция 160
Трансмембранные белки 223, 230
Транспортная РНК, *см.* тРНК
Трансферрин 33
Третичная структура 16—17, 35, 40, 70—73, 84, 96—103, 114—117, 121—122, 178
Триглицериды 231—232
Трипсин 60, 82
Трипсина ингибитор 73, 115, 282
Триптофан 113
тРНК 155, 160, 177—190, 311
— инициирующие 160, 162
— «клеверный лист» 179
— миграция ацила 155
— мини-петля 179
— неизменные остатки 180
— семейства (классы) 160—161
— D-спираль 179
— тирозиновый супрессор 161

тРНК, фотосшивки 187—188
— химическая модификация 186—188
$тРНК^{Phe}$ 180
Тройная (трехцепочечная) спираль 92, 93, 177—178, 198—199
Тропоколлаген 92
Тушение флуоресценции 113

Углеводы в белках 59. *См. также* Гликопротеиды
— составная часть мембран 221—222, 230
Углы внутреннего вращения 239
— поворота 253
Узнавание определенных участков нуклеиновых кислот 176
Ультрацентрифугирование 134—135
Уровни структурной организации 12, 41

Фазовое разделение в бислоях 220
Фазовые диаграммы 40, 144—145
Фазовый переход в липидах 221
Фенилаланиновая тРНК ($тРНК^{Phe}$) 186—187
Ферредоксин 74, 94, 96
Ферритин 29, 227—228
Фетуин 206
Фибриллярные белки 67, 72
Фибрин 68, 70
Филогенетическое древо глобинов 78
Флаводоксин 117, 120
Флуоресценции тушение 113
Формилметионин-тРНК ($тРНК_f^{Met}$) 162—163
Фосфоглицераткиназа 120
Фосфоглицериды 217
Фосфолипиды 214, 233
Фосфорилирование 60
Фотосшивки в тРНК 187—188

Хелатный эффект 27
Хиломикроны 231
Химическая модификация 70
Химотрипсин 16, 18, 30, 82—83, 109, 112, 120
Холестерин 217—218, 231—232
Холофермент 62
Хроматин 212

Цвиттерионы 43
Цинк (Zn^{2+}) 61, 102, 135
Цитохром *с* 221
Цитохромоксидаза 58

Четвертичная структура 19, 119, 129, 133—139, 192

Шаг спирали 14
Шпильки 156

Эволюция 28—29, 77, 120—121
Эластаза 83—84
Эластин 68—70
Электронная микроскопия 165—167, 224
Электронной плотности карта 88
Электронно-микроскопический маркер 165
Электростатическое взаимодействие 28, 44, 47—51
Электрофорез 135—137
Эндоплазматический ретикулум 234
Эритроцита мембрана 218, 229
Эстеразы сериновые 30
Эффекторы 36

ЯМР 113
Янтарный ангидрид 135, 137

Оглавление

Предисловие редакторов перевода 5
Предисловие 6

ТОМ 1

КОНФОРМАЦИЯ БИОЛОГИЧЕСКИХ МАКРОМОЛЕКУЛ

Глава 1. Стратегия и тактика биофизической химии. *Перевод А. В. Вологодского* 10

1.1. Уровни структурной организации биологических макромолекул 10

Первичная структура, т.е. последовательность звеньев 13
Вторичная структура отвечает спиралям, построенным из мономерных звеньев 14
Третичная структура, т.е. пространственное расположение звеньев 17
Четвертичная структура, т.е. расположение субъединиц 19

1.2. Некоторые центральные вопросы биофизической химии 22

Качество препаратов 22
Вопросы о структуре 24
Вопросы, касающиеся функции 32

1.3. Некоторые стратегические подходы, используемые в биофизической химии 37

Использование системы меньших размеров в качестве модели 38
Наблюдение лишь за частью системы 38
Сравнение между собой почти идентичных систем 38
Выделение дискретных состояний системы 40

Краткие выводы 41
Задачи 42
Литература 42

Глава 2. Структура белков. *Перевод Л. А. Белорусского* 43

2.1. Свойства аминокислот 43

Ионизация амино- и карбоксильных групп в аминокислотах и пептидах 43
Ионизация боковых групп 47
Ионизационное равновесие в белках 48
Полярность боковых групп аминокислот 52

2.2. Химический состав белков 54

Аминокислотный состав 54
Предсказание свойств белков по аминокислотному составу 57
Многие белки содержат необычные аминокислоты 58
Металлы и простетические группы 60

2.3. Первичная структура ... 65

Дисульфидные и другие поперечные связи ... 66
Первичная структура и анализ вторичной и третичной структуры ... 70
Первичная структура и предсказание вторичной и третичной структуры ... 70
Аминокислотная последовательность белка и анализ его функций ... 73
Сравнение аминокислотных последовательностей родственных белков ... 77
Мутантные белки; гемоглобин ... 79

2.4. Вторичная структура ... 87

α-Спираль ... 89
β-Слои и другие типы вторичной структуры ... 91
Полипролиновые спирали и коллаген ... 92
Распространенность вторичных структур в белках ... 94

2.5. Третичная структура ... 96

Общая организация пептидного скелета ... 103
Окружение отдельных остатков ... 106
Плотность упаковки аминокислотных остатков в свернутых молекулах белка ... 109
Объем и плотность белков ... 111
Динамична или статична третичная структура? ... 113
Методы сравнения третичных структур ... 114
Связь между структурным и функциональным подобием ... 119

2.6. Четвертичная структура ... 122

Соображения симметрии ... 122
Анализ числа субъединиц и полипептидных цепей ... 129
Взаимное расположение субъединиц ... 132
Четвертичная структура аспартат-транскарбамоилазы ... 133
Стабильность четвертичных глобулярных структур ... 138
Спиральные четвертичные структуры ... 139
Определение длины спиральных ассоциатов ... 143

Краткие выводы ... 146
Задачи ... 146
Литература ... 147

Глава 3. Структура нуклеиновых кислот. *Перевод А. В. Вологодского* ... 148

3.1. Свойства нуклеозидов и нуклеотидов ... 148

Ионизационное равновесие нуклеозидов и нуклеотидов ... 149

3.2. Нуклеотидный состав ДНК и РНК ... 153

Правила Чаргаффа для ДНК ... 153
Нуклеотидный состав РНК ... 154
Другие компоненты нуклеиновых кислот ... 155

3.3. Первичная структура ... 156

Строение ковалентной цепи ... 156
Нуклеотидные последовательности ... 157
Предположения о вторичной структуре, основанные на сравнении последовательностей тРНК ... 160
Информация о последовательности и анализ структуры и функции ... 162
Методы прямого сравнения последовательностей ... 164

3.4. Вторичная структура ... 167

Спаривание оснований — главная особенность двухцепочечных нуклеиновых кислот ... 168

Структура двойных спиралей нуклеиновых кислот 170
Структурные различия между РНК и ДНК 175
Другие полинуклеотидные спирали 177

3.5. Третичная структура 178

Третичные взаимодействия в тРНК с участием водородных связей 178
Стэкинг оснований в тРНК 186
Третичная структура $тРНК^{Phe}$; биохимические и химические свойства этой молекулы 186
Третичная структура ДНК 190

3.6. Четвертичная структура 192
Краткие выводы 192
Задачи 194
Литература 194

Глава 4. Другие биологические полимеры. *Перевод А. А. Полежаева* 195

4.1. Полисахариды 195

Первичная структура 195
Уровни структурной организации полисахаридов 198
О структуре и функциях полисахаридов 200

4.2. Ассоциаты, образованные из разных макромолекул 201

Комплексы сахаридов с пептидами, белками или липидами в клеточных стенках бактерий 201
Гликопротеиды на поверхности животных клеток 204
Нуклеопротеиды: нековалентные комплексы между белками и нуклеиновыми кислотами 207
Белки и нуклеиновые кислоты вирусов 209
Взаимодействие между РНК и белками в рибосомах 211
Хроматин: иерархия организованных структур 212

4.3. Липиды в биологических мембранах 213

Липидные компоненты мембран 214
Чистые липиды в бислоях 218
Смеси липидов в бислоях 219
Фазовые переходы в липидах 221

4.4. Белки в биологических мембранах 221

Белковые, липидные и углеводные компоненты 221
Типы и расположение белков 221
Поверхностная локализация белков 223
Латеральное расположение белков 224
Электронная микроскопия мембран, обработанных методом замораживания — травления 224
Распределение отдельных макромолекул, полученное с помощью электронного микроскопа 226

Общая картина строения биологических мембран 228
Мембрана эритроцита 229

4.5. Взаимодействия между белками и липидами 230

Взаимодействия между белками и липидами в бислое 230
Липопротеиды 231
Невыявленные макромолекулярные комплексы 234
Краткие выводы 234

Задачи ... 235
Литература ... 236

Глава 5. Конформационный анализ и силы, определяющие структуру белка. *Перевод Л. А. Белорусского* ... 237

5.1. Основные проблемы, связанные со структурой белковых молекул ... 237

5.2. Геометрия полипептидной цепи ... 237

Углы внутреннего вращения ... 239
Конформационные контурные карты, или диаграммы Рамачандрана ... 241

5.3. Оценки потенциальной энергии ... 243

Взаимодействия между валентно не связанными атомами ... 244
Дипольные взаимодействия ... 245
Собственные торсионные потенциалы, присущие одинарным связям ... 247
Общая энергия как сумма отдельных вкладов ... 247

5.4. Результаты расчетов потенциальной энергии ... 248

Остатки глицил и L-аланил ... 248
Роль дипольных взаимодействий ... 248
Выводы из анализа конформационных энергий глицила и L-аланила ... 249
Конформационная энергия некоторых других остатков ... 250
Изолированный остаток L-пролил ... 251
Конформационная энергия остатка, предшествующего L-пролилу ... 252
Конформационные ограничения, налагаемые на остаток, предшествующий пролилу ... 253

5.5. Экспериментально наблюдаемые значения углов поворота ... 253

Пример лизоцима ... 255

5.6. Водородные связи ... 257

Конкуренция молекул воды за места образования водородных связей в белке ... 258

5.7. Гидрофобные взаимодействия и структура воды ... 260

Физические свойства воды ... 260
Влияние растворителя на структуру белка ... 261
Неполярные молекулы в воде ... 262
Вклады аддитивной энтропии и энтропии смешения в свободную энергию переноса ... 263
Данные по свободной энергии переноса ... 265
Взаимодействие боковых групп неполярных аминокислот с водой ... 266
Понятие о гидрофобных взаимодействиях ... 267
Нарушение гидрофобных взаимодействий мочевиной ... 267

5.8. Ионные взаимодействия ... 268

Условия возникновения ионных взаимодействий в белках ... 268
Упрощенное рассмотрение ионных взаимодействий ... 269
Роль энтропии в образовании ионных пар ... 270
Сравнение гидрофобных и ионных взаимодействий ... 270

5.9. Дисульфидные связи ... 270

Возможные способы связывания N полуцистинов ... 271
Образование специфических дисульфидных связей ... 271
Реокисленные рибонуклеазы с образованием нативного белка ... 271
Реокисление проинсулина с образованием нативного инсулина ... 274

5.10. Данные о структуре белков 274
Число возможных конформаций 274
Сравнение структур лизоцима и миоглобина 275
5.11. Предсказание структуры белков 277
Методы предсказания структуры 277
Пример предсказания структуры 278
Сравнение предсказаниых и наблюдаемых вторичных структур 281
5.12. Построение молекулярных моделей с помощью ЭВМ 283
Краткие выводы 283
Задачи 285
Литература 285

Глава 6. Конформационный анализ и силы, определяющие структуру нуклеиновых кислот. *Перевод А. В. Вологодского* 287
6.1. Общая характеристика структуры нуклеиновых кислот 287
6.2. Геометрические аспекты 287
6.3. Поворотная изомерия относительно гликозидной связи; конформация рибозы 289
6.4. Углы вращения остова и стерические ограничения 289
Взаимодействия первого порядка 290
Взаимодействия второго порядка 292
Дальиейший анализ 292
6.5. Силы, стабилизирующие упорядоченные конформации 294
6.6. Спаривание оснований 295
Уотсон-криковское спаривание в кристаллах диинуклеотидов 297
Специфическое спаривание оснований в растворе 298
6.7. Стэкинг оснований 302
Стэкинг оснований в одноцепочечной poly (A) в растворе 304
Стэкинг мононуклеозидов в растворе 306
Пуриновый стэкинг сильнее, чем пиримидиновый 308
Роль растворителя в стабилизации стэкинга 309
6.8. Третичная структура нуклеиновых кислот 311
Краткие выводы 312
Задачи 312
Литература 313
Приложение А. Основы матричной алгебры 315
Приложение Б. Решение задач 319
Предметный указатель 323